

GenCore version 5.1.6  
(c) 1993 - 2005 Compugen Ltd.

Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw model

Run on: August 25, 2005, 03:18:03 ; Search time 166 Seconds  
(without alignments)

382.101 Million cell updates/sec

Title: US-10-664-038-1

Perfect score: 857

Sequence: 1 ARILNLKKKTNVTPTVAHRT.....ALPGPWRMASGFNKATAMQQ 164

Scoring table: BLOSUM62

dapox 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing First 45 summaries

Database : A\_Geneseq\_16Dec04:  
1: geneseqDP19308;\*  
2: geneseqDP19308;\*  
3: geneseqDP20018;\*  
4: geneseqDP20018;\*  
5: geneseqDP20028;\*  
6: geneseqDP20038;\*  
7: geneseqDP20038;\*  
8: geneseqDP20048;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	97.7	160	8 ADG22790	Adg22790 Hepatitis AAY44545 Hepatitis Abu61864 HCV core
2	813	94.9	198	3 AAY44545	Aay44545 Hepatitis Abu61864 HCV core
3	806	94.0	161	6 ABU61864	Abu61864 HCV core
4	768	89.6	161	6 ABU61866	Abu61866 HCV core
5	763	89.0	161	6 ABU61865	Abu61865 HCV core
6	645	75.3	198	3 AAY44558	Aay44558 Hepatitis Abu61870 HCV core
7	624	72.8	161	6 ABU61870	Abu61870 HCV core
8	583	68.0	198	3 AAY44554	Aay44554 Hepatitis Abu61868 HCV core
9	552	64.4	143	6 ABU61868	Abu61868 HCV core
10	548	63.9	143	6 ABU61872	Abu61872 HCV core
11	547	63.8	143	6 ABU61871	Abu61871 HCV core
12	540	63.0	143	6 ABU61867	Abu61867 HCV core
13	540	63.0	143	6 ABU61873	Abu61873 HCV core
14	540	63.0	143	6 ABU61874	Abu61874 HCV core
15	534	62.3	143	6 ABU61875	Abu61875 HCV core
16	508	59.3	198	3 AAY44557	Aay44557 Hepatitis Aay44557 Hepatitis
17	501	58.5	198	3 AAY44552	Aay44552 Hepatitis Abu61879 HCV core
18	487	56.8	154	6 ABU61879	Aay44556 Hepatitis Aay44556 Hepatitis
19	464	54.1	198	3 AAY44556	Aay44553 Hepatitis Abu61869 HCV core
20	441	51.5	198	3 AAY44553	Aay44553 Hepatitis Abu61873 HCV core
21	428	49.9	139	6 ABU61869	Aay44557 Hepatitis Aay44557 Hepatitis
22	426	49.7	99	8 ADR22073	Adr22073 Anti-hepatitis Adr21979 Anti-hepatitis
23	424	49.6	99	8 ADR21979	Adr21979 Anti-hepatitis Abu61878 HCV core
24	424	49.5	154	6 ABU61878	Abu61878 HCV core
25	422	49.2	99	8 ADR21949	Adr21949 Anti-hepatitis

## ALIGNMENTS

RESULT 1  
ID ADG22790 standard; protein; 160 AA.  
XX ADG22790;  
XX AC;  
XX DT 26-FEB-2004 (first entry)  
XX DE Hepatitis C virus F protein.  
XX KW virucide; antiviral RNA silencing pathway modulator; gene inactivation; RNA silencing; recombinant construct; antiviral RNA silencing pathway; viral infection; hepatitis C virus; F protein; RNA silencing suppressor.  
XX OS Hepatitis C virus.  
XX PN US2003219407-A1.  
XX XX  
XX PD 27-NOV-2003.  
XX PP 15-MAY-2002; 2002US-00150283.  
XX PR 15-MAY-2002; 2002US-00150283.  
XX PA (REGC ) UNIV CALIFORNIA.  
XX PI Ding S, Li H, Li W;  
XX DR WPI; 2004-021927/02.  
DR N-PSBB; ADG22791.  
XX PT New recombinant DNA construct comprising a polynucleotide sequence of a virus which activates RNA silencing, and a polynucleotide sequence of a gene, useful for inactivating a gene in a cell.  
XX PS Disclosure; SEQ ID NO 1; 20pp; English.  
XX CC The invention describes a recombinant DNA construct for inactivation of a gene in a cell. The construct comprises: a polynucleotide sequence of a virus which activates RNA silencing; and a polynucleotide sequence of the gene. The recombinant constructs are useful in inactivating a gene in a cell or in an animal. Modulators of antiviral RNA silencing pathway are useful for treating or preventing viral infection. This is the amino acid sequence 160 AA.;  
SQ Sequence 160 AA.;  
SQ Query Match 97.7%; Score 837; DB 8; Length 160;



CC be used in immunoassays for detecting HCV antigens and/or antibodies in  
 CC samples for the diagnosis of HCV infections. The present sequence  
 CC represents a p17 protein from an HCV strain.

XX Sequence 161 AA;

Query Match 94.0%; Score 806; DB 6; Length 161;  
 Best Local Similarity 99.4%; Pred. No. 7.5e-1;  
 Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 SQ

Qy 7 KCKKINVTPTVAHTSSRAVRSLEFTCRAGLDWTCARRRLPSGRNLEYDVSLSPR 66  
 Db 8 QKCTINVTPTVAHTSSRAVRSLEFTCRAGLDWTCARRRLPSGRNLEYDVSLSPR 67

Qy 67 LVGRBAGPGLSPGTLPGSNAMRAGGRGDSCLPVALGLAGAPOTPGVGRAIYRSSIPLR 126  
 Db 68 LVGRBAGPGLSPGTLPGSNAMRAGGRGDSCLPVALGLAGAPOTPGVGRAIYRSSIPLR 127

Qy 127 AASPTSWGTYRSSAPELLLEALPGPWRMASCFWKTA 160  
 Db 128 AASPTSWGTYRSSAPELLLEALPGPWRMASCFWKTA 161

RESULT 4  
 ABU61866 standard, protein; 161 AA.

XX ABU61866;  
 AC ABU61866;  
 XX 23-OCT-2003 (revised)  
 DT 14-AUG-2003 (First entry)

XX DE HCV core protein frameshift protein p17 #3.

XX KW hepatitis C infection; RNA frameshift; core protein; p17; viricide;

XX hepatotropic; overlapping open reading frame; p21c; vaccine.

XX OS Hepatitis C virus; genotype 1a.

XX US2002076415-A1.

XX PN 20-JUN-2002.

XX PP 14-DEC-2000; 2000US-00736959.

XX PR 14-DEC-1999; 99US-0170835P.

XX DR 2003-479366/45.

XX PS Claim 9; Page 14; 37pp; English.

XX The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping

CC open reading frame in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting

CC HCV infections.

XX PT Isolated hepatitis C virus (HCV) proteins formed by expression of the

PT overlapping open reading frames in the core protein gene sequence through

PT a frame shifting mechanism, useful for vaccinating against, and detecting

PT PT the invention relates to an isolated and purified protein of the

PT hepatitis C virus (HCV) that is formed by expression of an overlapping

PT open reading frame in the core protein gene sequence through an RNA frame

PT shifting mechanism. The protein is termed p17 (the full length, unshifted

PT protein being p21c). Also included are a vaccine (including a DNA

PT vaccine) for immunising a mammal against hepatitis C (producing a

PT protective antibody) comprising at least 1 protein of p17 (or a nucleic

PT acid encoding p17), an anti-viral composition (used to treat hepatitis C)

PT comprising a compound that binds to p17, antibodies directed against an

PT HCV core protein which are elicited by immunising an animal using the

PT partially purified protein p17, a method for analysing an HCV antigen in

PT a sample using the anti-p17 antibodies and detection of anti-HCV

PT antibodies in a sample using the p17 proteins. The HCV p17 and the DNA

PT sequences that encode it may be used as vaccines for immunising patients

PT against HCV infection. The antibodies and the antiviral compound may also

PT be used for treating HCV infections. HCV p17 and the antibodies may also

PT be used in immunoassays for detecting HCV antigens and/or antibodies in

PT samples for the diagnosis of HCV infections. The present sequence

PT represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to

PT standardise OS field)

XX SQ Sequence 161 AA;

Query Match 89.6%; Score 768; DB 6; Length 161;

Best Local Similarity 95.5%; Pred. No. 5.5e-70;

Matches 147; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

XX Qy 7 KCKKINVTPTVAHTSSRAVRSLEFTCRAGLDWTCARRRLPSGRNLEYDVSLSPR 66  
 Db 8 QKCTINVTPTVAHTSSRAVRSLEFTCRAGLDWTCARRRLPSGRNLEYDVSLSPR 67

Qy 67 LVGRBAGPGLSPGTLPGSNAMRAGGRGDSCLPVALGLAGAPOTPGVGRAIYRSSIPLR 126  
 Db 68 LVGRBAGPGLSPGTLPGSNAMRAGGRGDSCLPVALGLAGAPOTPGVGRAIYRSSIPLR 127

Qy 127 AASPTSWGTYRSSAPELLLEALPGPWRMASCFWKTA 160  
 Db 128 AASPTSWGTYRSSAPELLLEALPGPWRMASCFWKTA 161

RESULT 5

ABU61865

ID ABU61865 standard; protein; 161 AA.

XX AC ABU61865;

XX DT 23-OCT-2003 (revised)

XX DT 14-AUG-2003 (First entry)

XX DE HCV core protein frameshift protein p17 #2.

XX KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; viricide;

XX hepatotropic; overlapping open reading frame; p21c; vaccine.

XX OS Hepatitis C virus; genotype 1a.

XX PN US2002076415-A1.

XX PN 20-JUN-2002.

XX PR 14-DEC-2000; 2000US-00736959.

XX PR 14-DEC-1999; 99US-0170835P.

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/45.

XX PS (OUJJ/) OU J.

XX PA (XUZZ/) XU Z.

XX PI On J, Xu Z;

XX DR WPI; 2003-479366/

acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The Hcv p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the antiviral compound may also be used for treating HCV infections. Hcv p17 and the antiviral compound may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-Oct-2003 to standardise OS field)

XX

SQ Sequence 161 AA;

Query Match Score 763; DB 6; Length 161;  
Best Local Similarity 94.8%; Pred. No. 1.8e-69;  
Matches 146; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
Qy 7 KKKCNVTPVTAHRSSSRVAVRSLVEFTCRAGLDWVTCARRBLPSGRNLEVDVSISPR 66  
Db 8 QRKENVTPVTAHRSSSRVAVRSLVEFTCRAGLDWVTCARRBLPSGRNLEVDVSISPR 67  
Qy 67 LVGPRAGPGLSPGTLPGSAMRAAGRDGSCLPVALGLIAPOTPGVGRAIWRSSTPLR 126  
Db 68 HVGPRAGPGLSPGTLPGSAMRAAGRDGSCLPVALGLIAPOTPGVGRAIWRSSTPLR 127  
Qy 127 AASPTSWGTYRSSAPLLEALPGPKRMASGFWKTA 160  
Db 128 AASPTSWGTYRSSAPLLEALPGPKRMASGFWKTA 161

RESULT 6

AAY44558 ID AAY44558 standard; protein; 198 AA.  
XX AC AAY44558;  
XX DT 04-APR-2000 (first entry)

Hepatitis C virus protein encoded by DNA isolate D50482.  
KW Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine;  
KW prevention; diagnosis; therapeutic target; anti-HCV therapy;  
KW HCV infection; DNA isolate D50482.

XX OS Hepatitis C virus.

XX KEY Location/Qualifiers  
PT Misc-difference 139  
FT /label= unknown  
FT /note= "Corresponds to stop codon"  
PT Misc-difference 143  
FT /label= unknown  
FT /note= "Corresponds to stop codon"  
PT Misc-difference 161  
FT /label= unknown  
FT /note= "Corresponds to stop codon"  
PT Misc-difference 178  
FT /label= unknown  
FT /note= "Corresponds to stop codon"  
PT Misc-difference 184  
FT /label= unknown  
FT /note= "Corresponds to stop codon"

XX PN W09963941-A2.

XX XX 16-DEC-1999.

XX PP 09-JUN-1999; 99WO-US012929.

XX PR 09-JUN-1998; 98US-00886705.

PR 11-JUN-1998; 98US-0089138P.

XX PA (BRAN/) BRANCH A.D.

PA (WALE/) WALEWSKI J.L.  
(STUM/) STUMP D.D.

XX PI Branch AD, Walewski JL,

XX DR WPL; 2000-126431/11.

Novel Hepatitis C virus peptides useful in vaccine compositions, for diagnosing HCV infection and as therapeutic agents.

XX PS Example 1; Page 39-41; 50PP; English.

XX

The present sequence is a Hepatitis C virus (HCV) protein derived from an HCV isolate AAB50482. The novel HCV protein is not encoded by the standard HCV open reading frame but by an alternate reading frame (ARF) which is +1 or +2 to the standard HCV ORF. The position of the first nucleotide of the ARF may vary slightly depending upon the isolate. The protein elicits an immune response in patients infected with HCV and are produced during HCV infection. The present sequence is used in vaccine compositions for preventing HCV infection. It is also used for diagnosing HCV infection and as a target for anti-HCV therapy

XX Sequence 198 AA;

Query Match Score 75.3%; DB 3; Best Local Similarity 76.4%; Pred. No. 2.2e-57; Matches 126; Conservative 8; Mismatches 30; Indels 1; Gaps 1;

Qy 1 ARTLNKXKKTKTNTPTVTAHRTSSSRVAVRSLVEFTCCRAGLDWVCARRRLPSGRNLVEVD 60

Db 1 AQILNLCKPKNTPTVTAHRTSSSRVAVRSLVEFTCCRAGLDWVCARRRLPSGRNLVEVG 60

Qy 61 VSLSPRLVGPRAGPGLSPGTLPGSMSMARAAGRDGSCLPVALGLASAPOTPGVRAIWR 120

Db 61 DNLSPRFAGPRAGPGLSPGTLPGSMSMAMRVNGQDGSCHPGALGVGAAPRTPGVRAIWR 120

Qy 121 SSIPPLRASPTSWGTYRSSAPLLEALPGPKRMASGFWKTA-TMQQ 164

Db 121 SSIPPLRASPTSWGTYRSSAPLLEALPGPKRMASGFWKTA-TMQQ 165

RESULT 7  
ID ABU61870 standard; protein; 161 AA.

XX XX ABU61870;

AC AC

DT 23-OCT-2003 (revised)

XX DT 14-AUG-2003 (first entry)

DE DE

HCV core protein frameshift protein p17 #7.

XX KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;

KW KN hepatotropic; overlapping open reading frame; p21C; vaccine.

XX OS US2002076415-A1.

XX PN Hepatitis C virus; genotype 1b.

XX PN US2002076415-A1.

XX PD 20-JUN-2002.

XX PF 14-DEC-2000; 2000US-00736959.

XX PR 14-DEC-1999; 99US-017035P.

XX PA (OUJ/) OU J.

XX PA (XUZ/) XU Z.

XX PI Ou J, Xu Z;

XX

DR WPI; 2003-479366/45.  
 PT Isolated hepatitis C virus (HCV) proteins formed by expression of  
 PT overlapping open reading frames in the core protein gene sequence through  
 PT a frame shifting mechanism, useful for vaccinating against, and detecting  
 XX HCV infections.  
 PS Claim 9; Page 15; 37pp; English.

XX The invention relates to an isolated and purified protein of the  
 CC hepatitis C virus (HCV) that is formed by expression of an overlapping  
 CC open reading frame in the core protein gene sequence through an RNA frame  
 CC shifting mechanism. The protein is termed p17 (the full length, unshifted  
 CC protein being p21c). Also included are a vaccine (including a DNA  
 CC vaccine) for immunising a mammal against hepatitis C (producing a  
 protective antibody) comprising at least 1 protein of p17 (or a nucleic  
 acid encoding p17), an anti-viral composition (used to treat hepatitis C)  
 comprising a compound that binds to p17, antibodies directed against an  
 HCV core protein which are elicited by immunising an animal using the  
 partially purified protein p17, a method for analysing an HCV antigen in  
 a sample using the anti-p17 antibodies and detection of anti-HCV  
 antibodies in a sample using the p17 proteins. The HCV p17 and the DNA  
 sequences that encode it may be used as vaccines for immunising patients  
 against HCV infection. The antibodies and the antiviral compound may also  
 be used for treating HCV infections. HCV p17 and the antibodies may also  
 be used in immunoassays for detecting HCV antigens and/or antibodies in  
 samples for the diagnosis of HCV infections. The present sequence  
 represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to  
 CC standardise OS field)

XX Sequence 161 AA;

Query Match 72.8%; Score 624; DB 6; Length 161;  
 Best Local Similarity 77.3%; Pred. No. 2.6e-55;  
 Matches 119; Conservative 7; Mismatches 28; Indels 0; Gaps 0;

Qy 7 KKCCTNVTPTVAHRTSSRVAVRSLLVETCCTRAGALDWYCARRELPSGRNLVEYDVSLSPR 66  
 Db 8 QRKPNVTPTAHRTSSRVAVRSLLVETCCTRAGAGWVARLGRLPSGRNLVEGDNLSPR 67

Qy 67 LVGPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPYAHLGAGAQPTPGVGRAIWYRSSIPR 126  
 Db 68 LADPRAGPGLSPGTLGPSPMAMRAALGGDOSCHPAAPGLVGAAPTPGVGRIVIWRSSIPSH 127

Qy 127 AASPTSWGTYSRSSLAPPLLEALPGPWRMASGEFWKTA 160  
 Db 128 AASPTSWGTYSRSSLAPPLLEALPGPWRMASGEFWKTA 161

RESULT 8  
 AAY44554 standard; protein; 198 AA.  
 XX AAY44554 ;  
 AC DT 04-APR-2000 ; (first entry)  
 XX Hepatitis C virus protein encoded by DNA isolate D14853.  
 KW Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine;  
 KW prevention; diagnosis; therapeutic target; anti-HCV therapy;  
 KW HCV infection; DNA isolate D14853.  
 XX Hepatitis C virus  
 XX Key Location/Qualifiers  
 KW Misc-difference 143  
 FT /label= "Corresponds to stop codon"  
 FT Misc-difference 161  
 FT /label= unknown  
 FT Misc-difference 184  
 FT /note= "Corresponds to stop codon"  
 XX /label= unknown  
 /note= "Corresponds to stop codon"

PT PT /label= unknown  
 PT XX WO9633941-A2.  
 PT PN 16-DBC-1999.  
 PD XX PP 09-JUN-1999; 99WO-US012929.  
 PR XX PR 09-JUN-1998; 98US-0088670P.  
 PA XX PA 11-JUN-1998; 98US-0089138P.  
 PA (BRAN/ ) BRANCH A D.  
 PA (WALE/ ) WALEWSKI J L.  
 PA (STUM/ ) STUMP D D.  
 PI XX Branch AD, Walewski JL, Stump DD;  
 DR WPI; 2000-126431/11.

PT Novel Hepatitis C virus peptides useful in vaccine compositions, for  
 PT diagnosing HCV infection and as therapeutic agents.  
 PS Example 1; Page 39-41; 50pp; English.

CC The present sequence is a Hepatitis C virus (HCV) protein derived from an  
 CC HCV isolate AAD14853. The novel HCV protein is not encoded by the  
 CC standard HCV open reading frame but by an alternate reading frame (ARF)  
 CC which is +1 or +2 to the standard HCV ORF. The position of the first  
 CC nucleotide of the ARF may vary slightly depending upon the isolate. The  
 CC protein elicits an immune response in patients infected with HCV and are  
 CC produced during HCV infection. The present sequence is used in vaccine  
 CC compositions for preventing HCV infection. It is also used for diagnosing  
 CC HCV infection and as a target for anti-HCV therapy  
 SQ Sequence 198 AA;

Query Match 68.0%; Score 583; DB 3; Length 198;  
 Best Local Similarity 72.6%; Pred. No. 5.1e-51;  
 Matches 119; Conservative 3; Mismatches 42; Indels 0; Gaps 0;

Qy 1 ARTINLKEKTNVTPTVAHRTSSRVAVRSLLVETCCTRAGALDWYCARRELPSGRNLVEYD 60  
 Db 1 ARTINLKEKPNVTPTAHRTSSRVAVRSLLVETCCTRAGAPENVCARGLPSGRNLAGG 60

Qy 61 VSLRPRPLVGPRAGPGLSPGTLGPSPMAMRAAGGRDGSCLPVAVLGLAGAQPTPGVGRATWVR 120  
 Db 61 VSLRPRPAPDPREPGRSPTGLGPSPMTRAVGRDPSCEPPAAIGLVGALLTPGGGHAIWVR 120

Qy 121 SSTPLRASPTSWGTYSRSSLAPPLLEALPGPWRMASGEFWKTA 164  
 Db 121 SSIPSRVASPTSWGTYSRSSLAPPLLEALPGPWRMASGEFWKTAITQ 164

RESULT 9  
 ABU61868  
 ID ABU61868 standard; protein; 143 AA.  
 AC AC ABU61868;  
 XX AC ABU61868;  
 XX DT 23-OCT-2003 (revised)  
 DT 14-AUG-2003 (first entry)  
 XX DE HCV core protein frameshift protein p17 #5.  
 XX KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;  
 KW hepatotropic; overlapping open reading frame; p21c; vaccine.  
 OS OS Hepatitis C virus; genotype 1b.  
 XX FN Key Location/Qualifiers  
 KW Misc-difference 143  
 FT /label= unknown  
 FT Misc-difference 161  
 FT /label= unknown  
 FT Misc-difference 184  
 FT /note= "Corresponds to stop codon"

OS ID US2002076415-A1.  
 XX PD 20-JUN-2002.

XX	14-DEC-2000;	2000US-00736959.	OS	Hepatitis C virus; genotype 1b.
XX	PP	US2002076415-A1.	XX	
PR	14-DEC-1999;	99US-0170835P.	XX	
XX	PD	20-JUN-2002.	XX	
PA	(OUUJ/)	OU J.	XX	
PA	(XUZZ/)	XU Z.	XX	
PI	Ou J,	Xu Z;	XX	
XX	DR	WPI; 2003-479366/45.	XX	
XX	PT	Isolated hepatitis C virus (HCV) proteins formed by expression of overlapping open reading frames in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.	PT	Isolated hepatitis C virus (HCV) proteins formed by expression of overlapping open reading frames in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.
PS	Claim 9; Page 14;	37pp; English.	PS	Claim 9; Page 15; 37pp; English.
XX	The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21c). Also included are a vaccine (including a DNA acid encoding p17), an anti-viral composition (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the antiviral compound may also be used for treating HCV infections. HCV p17 and the antibodies may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to standardise OS field)	XX	The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21c). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the antiviral compound may also be used for treating HCV infections. HCV p17 and the antibodies may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to standardise OS field)	
XX	Sequence 143 AA;	Score 552; DB 6; Length 143;	SQ	Sequence 143 AA;
Qy	Best Local Similarity 64.4%; Matches 107;	Pred. No. 4.9e-48; Conservative 79.3%; Mismatches 5;	Query Match	Score 548; DB 6; Length 143;
Db	Indels 0; Gaps 0;	Best Local Similarity 78.1%; Matches 106;	Best Local Similarity 78.1%; Matches 106; Conservative 6;	Pred. No. 1.3e-47; Indels 0; Gaps 0;
Qy	Score 552; DB 6; Length 143;	Score 548; DB 6; Length 143;	Qy	7 KCKTNVTPTAHRTSSRVAVSLVEFTCRAAGLDIVCARRRLPSGRNLEYDVSLSPR 66
Db	DB 6; Mismatches 5;	DB 6; Mismatches 6;	Db	8 QRKPNVTLTAHRTSSRVAVSLVEFTCRAAGPDGVTCARLGRPLSPRNLYVGDNLSPR 67
Qy	DB 6; Mismatches 23;	DB 6; Mismatches 23;	Qy	67 LVGPRAGPGLSPGTIGPSAMRAAGGRDGSCLPVALGLAGAPQPVGRAIWRSIPLR 126
Db	DB 6; Mismatches 127;	DB 6; Mismatches 127;	Db	68 LAGPRAGPGLSPGTIGPSAMRAAGGDGSCHPAGLVLVGAAPMTPGVGRIVWTRSSIPSH 127
Qy	DB 6; Mismatches 127;	DB 6; Mismatches 127;	Qy	127 AASPTSMWGTYRSSAP 141
Db	DB 6; Mismatches 128;	DB 6; Mismatches 128;	Db	128 AASPTSMWGTYRSSAP 142
XX	AC ABU61872;	AC ABU61872;	RESULT 1.0	127 AASPTSMWGTYRSSAP 141
XX	DT 23-OCT-2003 (revised)	DT 23-OCT-2003 (revised)	ID ABU61871 standard; protein; 143 AA.	DB ABU61871;
XX	DT 14-AUG-2003 (first entry)	DT 14-AUG-2003 (first entry)	ID ABU61871;	XX ABU61871;
DB	HCV core protein frameshift protein p17 #9.	DB 23-Oct-2003 (revised)	DT 23-Oct-2003 (first entry)	XX ABU61872
XX	HCV; hepatitis C infection; RNA frame shift; core protein; p17; viricide; hepatotropic; overlapping open reading frame; p21c; vaccine.	XX ABU61872	DB 128 AASPTSMWGTYRSSAP 142	XX ABU61872

DE	HCV core protein frameshift protein p17 #8.	AC	ABU61867;
XX	HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide; hepatotropic; overlapping open reading frame; p21C; vaccine.	XX	23-OCT-2003 (revised)
KW	hepatotropic; overlapping open reading frame; p21C; vaccine.	DT	14-AUG-2003 (first entry)
XX		XX	
OS	Hepatitis C virus; genotype 1b.	DE	HCV core protein frameshift protein p17 #4.
XX		XX	
PN	US2002076415-A1.	KW	HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide; hepatotropic; overlapping open reading frame; p21C; vaccine.
XX	PD 20-JUN-2002.	XX	
XX	PP 14-DEC-2000; 20000US-00736959.	OS	Hepatitis C virus; genotype 1b.
XX	PR 14-DEC-1999; 99US-0170835P.	Key	Location/Qualifiers
XX	PA (OUJJ/) OU J. (XUZZ/) XU Z.	FT Misc-difference 70 /label= UNKNOWN /note= "Any amino acid"	
PI	Ou J., Xu Z;	FT	
XX	DR WPI; 2003-479366/45.	XX	
XX	PT Isolated hepatitis C virus (HCV) proteins formed by expression of overlapping open reading frames in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.	PP 14-DEC-2000; 20000US-00736959.	
PT		XX	
PT		PR 14-DEC-1999; 99US-0170835P.	
PT		XX	
PS	Ou J., Xu Z;	PA (OUJJ/) OU J. (XUZZ/) XU Z.	
XX	DR 20-JUN-2002.	PA	
XX	XX	XX	
CC	Claim 9; Page 15; 37pp; English.	PI	Ou J., Xu Z;
CC	The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21C). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequences that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the antiviral compound may also be used for treating HCV infections. HCV p17 and the antiviral compound may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to standardise OS field)	XX	XX
CC	Sequence 143 AA;	XX	XX
CC	Query Match 63.8%; Score 547; DB 6; Length 143;	XX	XX
CC	Best Local Similarity 78.5%; Prod. No. 1.e-47;	XX	XX
CC	Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;	XX	XX
Qy	7 KCKKNTVPTVAHRTSSSERAVRSLSLVEFTCCRAGALDWCARRRLPSGRNLIEVDVSLSLSPR 66	7 KCKKNTVPTVAHRTSSSERAVRSLSLVEFTCCRAGALDWCARRRLPSGRNLIEVDVSLSLSPR 66	7 KCKKNTVPTVAHRTSSSERAVRSLSLVEFTCCRAGALDWCARRRLPSGRNLIEVDVSLSLSPR 66
Db	8 QRKPNVTPTAAHTTSSRAVRSLSLVEFTCCRAGAPGTVCARGLPGRNLYVGDNLISPR 67	8 QRKPNVTPTAAHTTSSRAVRSLSLVEFTCCRAGAPGTVCARGLPGRNLYVGDNLISPR 67	8 QRKPNVTPTAAHRTSSSERAVRSLSLVEFTCCRAGAPGTVCARGLPGRNLYVGDNLISPR 67
Qy	67 LYOPRAGEFGLSPGTGLGPSMAMBAAGGRGDSCLPVALGLAGAPOPFGYGRAIWRSSSPLR 126	XX	XX
Db	68 LASPRAGEFGLSPGTGLGPSMAMBAAGGRGDSCLPVALGLAGAPOPFGYGRAIWRSSSPLR 127	SQ Sequence 143 AA;	XX
Qy	127 AASPTNSWGTYRSSAP 141	Query Match 63.0%; Score 540; DB 6; Length 143;	XX
Db	128 AASPTNSWGTFRSSAP 142	Best Local Similarity 77.8%; Prod. No. 8.2e-47;	XX
Qy	67 LVGPRAAGPGLSPGTGLGPSMAMRAAGGRGDSCLPVALGLAGAPOPFGYGRAIWRSSSPLR 126	Matches 105; Conservative 6; Mismatches 24; Indels 0; Gaps 0;	XX
Db	68 LAXPRAAGPGLSPGTGLGPSMAMRANGQDGSCHPAAPGLVGAAPRTGVRVWWNSSPLR 127	8 QRKPNVTPTAAHRTSSSERAVRSLSLVEFTCCRAGAPGTVCARGLPGRNLYVGDNLISPR 67	RESULT 12
XX	ID ABU61867	Qy 67 LVGPRAAGPGLSPGTGLGPSMAMRAAGGRGDSCLPVALGLAGAPOPFGYGRAIWRSSSPLR 126	
XX	RESULT 12	Db 68 LAXPRAAGPGLSPGTGLGPSMAMRANGQDGSCHPAAPGLVGAAPRTGVRVWWNSSPLR 127	



Best Local Similarity 77.8%; Pred. No. 8.2e-47; Matches 105; Conservative 5; Mismatches 25; Indels 0; Gaps 0;

CC standardise OS field  
XX SQ Sequence 143 AA:

Query Match 62.3%; Score 534; DB 6; Length 143;  
Best Local Similarity 77.0%; Pred. No. 3.4e-46;  
Matches 104; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

Qy 7 KKCTNVTPTPAHRTSSRVAWSLVEFTCCRAGALDWVCAARRLPSGNLEVDVSLSPR 66  
Db 8 QRKPNVTPTPAHRTLSSRVAWSLVEFTCCRAGALDWVCAARRLPSGNLVEGDLSLR 67

Qy 67 LVDPRAAGPLSPGTIGPSMAMMRAAGGRDSCSCLPVALGLAGAPQTPGVRAIWRSSIPR 126  
Db 68 LAGPAGPLSPGTIGPSMAMMRAANGQDSSCHCPAAPGLVGAAPRTPGVSRIWTRSSTFSH 127

Qy 127 AASPTSWGTYRSSAP 141  
Db 128 AASPTSWGTYRSSAP 142

Qy 7 KKCTNVTPTVAHRTSSRVAWSLVEFTCCRAGALDWVCAARRLPSGNLEVDVSLSPR 66  
Db 8 QRKPNVTPTAAHTSSRVAWSLVEFTFRAGAPGWCAQRGLPSRNLYEGDNUSPR 67

Qy 67 LVGPRAGPGLSPGTIGPSMAMMRAAGGRDGSCLPVALGLAGAPQTPGVRAIWRSSIPR 126  
Db 68 LASPRAGPGLSPGTIGPSMAMMRAVGQGSCSHPVAPGLVGAQRTPGVRVWRSIFSH 127

RESULT 15

ABU61875 ABU61875 standard; protein; 143 AA.

XX AC ABU61875;

XX DT 23-OCT-2003 (revised)  
DT 14-AUG-2003 (first entry)

XX DE HCV core protein frameshift protein p17 #12.

XX KW hepatitis C infection; RNA frameshift; core protein; p17; virucide;  
KW hepatotropic; overlapping open reading frame; p21c; vaccine.

XX OS Hepatitis C virus, genotype 1b.

PN US2002076415-A1.

XX DR 2003-479366/45.

XX PD 20-JUN-2002.

XX PP 14-DEC-2000; 2000US-00736959.

XX PR 14-DEC-1999; 99US-0170835D.

XX PA (OUJJ/) OU J.  
PA (XUZZ/) XU Z.

XX PI Ou J, Xu Z;

XX PS Page 15; 37pp; English.

XX PT Isolated hepatitis C virus (HCV) proteins formed by expression of overlapping open reading frames in the core protein gene sequence through a frame shifting mechanism, useful for vaccinating against, and detecting HCV infections.

XX CC The invention relates to an isolated and purified protein of the hepatitis C virus (HCV) that is formed by expression of an overlapping open reading frame in the core protein gene sequence through an RNA frame shifting mechanism. The protein is termed p17 (the full length, unshifted protein being p21c). Also included are a vaccine (including a DNA vaccine) for immunising a mammal against hepatitis C (producing a protective antibody) comprising at least 1 protein of p17 (or a nucleic acid encoding p17), an anti-viral composition (used to treat hepatitis C) comprising a compound that binds to p17, antibodies directed against an HCV core protein which are elicited by immunising an animal using the partially purified protein p17, a method for analysing an HCV antigen in a sample using the anti-p17 antibodies and detection of anti-HCV antibodies in a sample using the p17 proteins. The HCV p17 and the DNA sequence that encode it may be used as vaccines for immunising patients against HCV infection. The antibodies and the antiviral compound may also be used for treating HCV infections. HCV p17 and the antibodies may also be used in immunoassays for detecting HCV antigens and/or antibodies in samples for the diagnosis of HCV infections. The present sequence represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 03:26:38 ; Search time 40 Seconds  
(without alignments)

394.489 Million cell updates/sec

Title: US-10-664-038-1  
Perfect score: 857  
Sequence: 1 ARILNKKKTNVTPVAHRT.....ALPGPWRMAGSFWKATAMQQ 164

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 summaries

## Database :

- 1: PIR\_79;\*
- 2: piz2;\*
- 3: piz3;\*
- 4: piz4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	374	43.6	95	2	S44213		core protein - hep hypothetical prote transcription fact transcription fact hypothetical prote ankyrin related pr transcription fact ABC transporter, A transcription fact transcription fact hypothetical prote protein F5J5.15 [1 collagen dpy-2 - C serum response fac immediate-early pr collagen alpha 1(I) hypothetical prote hypothetical prote period protein. Per period protein. Per dopamine receptor hypothetical prote elastin precursor hypothetical prote hypothetical prote hypothetical prote DNA (cytosine-5-) - galactokinase (EC collagen alpha 1(I transcription acti
2	97	11.3	363	2	S22542		
3	96	11.2	479	1	S22543		
4	96	11.2	485	1	S22543		
5	93.5	10.9	105	2	H72733		
6	90	10.5	1016	2	T19006		
7	89	10.4	479	1	A31753		
8	89	10.4	552	2	F5J311		
9	88.5	10.3	583	1	S22544		
10	87.5	10.2	478	1	I47154		
11	87	10.2	756	2	B72622		
12	86	10.0	1617	2	B86482		
13	85	9.9	360	2	T37285		
14	85	9.9	365	2	A39481		
15	85	9.9	825	1	EDEBXD		
16	85	9.9	1042	1	CGCH1S		
17	84	9.8	143	2	B72649		
18	84	9.8	152	2	T34642		
19	84	9.8	1113	2	T14260		
20	84	9.8	1115	2	T13955		
21	83.5	9.7	387	1	DYHUD4		
22	83	9.7	299	2	C70753		
23	83	9.7	784	2	A26601		
24	82.5	9.6	419	2	G70602		
25	82.5	9.6	538	2	T27156		
26	82.5	9.6	1573	2	S01845		
27	81.5	9.5	395	1	KISMG		
28	81.5	9.5	671	1	CGRTIS		
29	81.5	9.5	1442	2	T42607		

## ALIGNMENTS

hypothetical prote transcription fact probable MFS trans collagen alpha 1(I) collagen alpha 6(I) probable DNA-direc transcription fact transcription fact transcription fact transcription fact hypothetical prote hook-containing pr cytokine receptor probable infB - My probable PPE prote probable relaxase

hypothetical prote core protein - hepatitis C virus hepatitis C virus

C;Species: C;Accession: S44213;C;Revision: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

R;Peuchtt, H.H.

submitted to the EMBL Data Library, April 1994

A;Reference number: S44213

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-95 <PEU>

A;Cross-references: UNIPROT:Q68874; EMBL:X78950; NID:9475172; PIDN:CAA55547.1; PID:9860

RESULT 1

S44213

core protein - hepatitis C virus

C;Species: hepatitis C virus

C;Accession: S44213;C;Revision: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

Query Match 43.6%; Score 374; DB 2; Length 95;

Best Local Similarity 78.7%; Pred. No. 1. 6e-25;

Matches 74; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

Qy 11 NTPTVAHRTSSSRVAYRSLVEFTCRAGALDWVCAERRRLPSGRNLLEVDSLSPRLVGP 70

Db 2 NTPTAAVRTSSRAVRSLVEFTCRAGAPGNVCAERGLPLPSGRNLVGDNLSPRLASP 61

Qy 71 RAGPGLSGTGLPSMAVRWGGDGSCHPTAPGL 104

Db 62 RAGPGLSGTGLPSMAVRWGGDGSCHPTAPGL 95

RESULT 2

T34931

hypothetical protein SC3F9.09 SC3F9.09 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Dte: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C;Accession: T34931

R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1998

A;Reference number: 221562

A;Accession: T34931

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-363 <SEB>

A;Cross-references: UNIPROT:O69949; EMBL:AL023862; PIDN:CAA19632.1; GSPDB:GN00070; SCOEL

C;Genetics:

A;Gene: SCOEDB:SC3F9.09

Query Match 11.3%; Score 97; DB 2; Length 363;

Best Local Similarity 30.6%; Pred. No. 0. 4e-47;

Matches 44; Conservative 14; Mismatches 40; Indels 46; Gaps 8;

Qy 39 GALDWVCAERGLPSGRNLLEVDSLSPRLVGPSPRAGLPLPSGRNLVGDNLSPRLASP 111

**RESULT 3**  
**S22542**  
 transcription factor Oct-2, splice form Oct-2.2 - mouse  
 N;Alternative names: NF-A2; OTF-2  
 C;Species: Mus musculus (house mouse)  
 C;Accession: S22542; S50017; S47215  
 R;Wirth, T.; Priess, A.; Annweiler, A.; Zwilling, S.; Oeler, B.  
 Nucleic Acids Res. 19, 43-51, 1991  
 A;Title: Multiple Oct-2 isoforms are generated by alternative splicing.  
 A;Reference number: S22539; MUID:91187647; PMID:2011512  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-479 <WIR>  
 A;Cross-references: UNIPROT:Q00196; EMBL:X57937; NID:953483; PIDN:CAA41005.1; PID:953484  
 R;Matsuo, K.; Clay, O.; Kuenzler, P.; Georgiev, O.; Urbanek, P.; Schaffner, W.  
 Biol. Chem. Hoppe-Seyler 375, 675-683, 1994  
 A;Title: Short introns interrupting the Oct-2 POU domain may prevent recombination between  
 A;Reference number: S50016; MUID:95194574; PMID:7888800  
 A;Accession: S50017  
 A;Status: translation not shown  
 A;Molecule type: DNA  
 A;Residues: 168-377 <MAT>  
 A;Cross-references: EMBL:X81031; NID:9534039; PIDN:CAA56934.1; PID:9536781  
 C;Comment: This protein is a tissue-specific transcriptional trans-activator.  
 C;Genetics:  
 A;Gene: Oct-2  
 A;Map position: 7  
 A;Introns: 221/3; 269/1; 318/3  
 C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology  
 C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulator  
 P;202-269/Domain: POU domain homology <POU>  
 P;298-354/Domain: homeobox homology <HOX>

Query Match 11.2%; Score 96; DB 1; Length 479;  
 Best Local Similarity 25.4%; Pred. No. 0.76; Gaps 6;  
 Matches 34; Conservative 8; Mismatches 28; Indels 64; Gaps 6;

Qy 43 WVCARRR-----LPS-GRNLLEVDSLSPRLVGPRAAGPGLSP-----78  
 Db 344 WECNRQKEKRINPCSAAPMLPSPEK----PTSTSPHLVTPQGAGTLPLSQASSSLST 399

Qy 79 -----GTLGPNSHAMRAAGGRDGSCLPV-----ALGL 104  
 Db 400 VTTLSAVGTIHLPSRTAGGGGGGRALLPLNSIPSVTPPPATTNSTNSPQSSHSAAGL 459

Qy 105 AGAPQTPGVGRAIW 118  
 Db 460 SGL--NPSAGPGLW 471

**RESULT 4**  
**S22543**  
 transcription factor Oct-2 splice form Oct-2.3 - mouse  
 N;Alternative names: NF-A2 protein; OTF-2 protein  
 N;Contains: transcription factor Oct-2 splice form Oct-2.1; transcription factor Oct-2  
 C;Species: Mus musculus (house mouse)  
 C;Accession: S22543; S22541; S22540; JH0596; I46698; S47215  
 R;Wirth, T.; Priess, A.; Annweiler, A.; Zwilling, S.; Oeler, B.  
 Nucleic Acids Res. 19, 43-51, 1991

**RESULT 5**  
**H72733**  
 hypothetical protein APE0408 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaiava, H.; Takamiya, M.; Maeuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999

A;Title: Multiple Oct2 isoforms are generated by alternative splicing.  
 A;Reference number: S22539; MUID:91187647; PMID:2011512  
 A;Accession: S22543  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-62-85 <WIR>  
 A;Cross-references: EMBL:X57936; NID:953481; PIDN:CAA41004.1; PID:953482  
 A;Note: Splice form Oct-2.1  
 A;Accession: S22540  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-62-124-485 <WI4>  
 A;Cross-references: EMBL:X57941; NID:953491; PIDN:CAA41009.1; PID:953492  
 A;Note: Splice form Oct-2.6  
 R;Stoykova, A.S.; Steiner, S.; Erselius, J.R.; Hartzopoulos, A.K.; Grubbs, P.  
 Neuron 8, 541-558, 1992  
 A;Title: Mini-Oct and Oct-2c: two novel, functionally diverse murine Oct-2 gene product.  
 A;Reference number: JH0596; MUID:92198662; PMID:1550677  
 A;Accession: JH0596  
 A;Molecule type: mRNA  
 A;Residues: 1-62-85-130, 'A', 432, 'P', 434-473 <STO>  
 A;Experimental source: brain  
 A;Note: Splice form Oct-2c  
 R;Stepchenko, A.G.  
 Dokl. Akad. Nauk SSR 325, 175-178, 1992  
 A;Title: [Interaction of Oct-binding transcription factors with a large series of 'nonc  
 A;Reference number: I48698  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: mRNA  
 A;Residues: 'P'P, 189-430, 'A', 432, 'P', 434-464, 'W', 466-485 <REB>  
 A;Cross-references: EMBL:Y57089; NID:953493; PIDN:CAA40369.1; PID:953494  
 C;Comment: This protein is a tissue-specific transcriptional trans-activator.  
 A;Gene: Oct-2  
 A;Map position: I48698  
 A;M;Position: 7  
 A;Query Match 11.2%; Score 96; DB 1; Length 485;  
 Best Local Similarity 25.4%; Pred. No. 0.77;  
 Matches 34; Conservative 8; Mismatches 28; Indels 64; Gaps 6;  
 C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulator;  
 P;208-275/Domain: POU domain homology <POU>  
 P;304-360/Domain: homeobox homology <HOX>

Query Match 11.2%; Score 96; DB 1; Length 485;  
 Best Local Similarity 25.4%; Pred. No. 0.77;  
 Matches 34; Conservative 8; Mismatches 28; Indels 64; Gaps 6;

Qy 43 WVCARRR-----LPS-GRNLLEVDSLSPRLVGPRAAGPGLSP-----78  
 Db 350 WFPNRQKEKRINPCSAAPMLPSPK----PTSYSPHLVTPQGAGTLPLSQASSSLSTT 405

Qy 79 -----GTLGPNSHAMRAAGGRDGSCLPV-----ALGL 104  
 Db 406 VTLSSAVGTIHLPSRTAGGGGGGRALLPLNSIPSVTPPPATTNSTNSPQSSHSAAGL 465

Qy 105 AGAPQTPGVGRAIW 118  
 Db 466 SGL--NPSAGPGLW 477

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: H72733  
A;Status: preliminary  
A;Cross-references: UNIPROT:Q9YF32; DDBJ:AP000059; NID:95103911; PIDN:BAA79364\_1; PID:di  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE0408

Query Match 10.9%; Score 93.5; DB 2; Length 105;  
Best Local Similarity 33.7%; Pred. No. 0.27;  
Matches 34; Conservative 11; Mismatches 33; Indels 23; Gaps 6;

Qy 64 SPRLVG--PRAAGPLSPGTLGPSPNAMRAAG-----GRDGSCLPVVA---LGLAGA 107  
Db 5 STRIVSGTULPRASPV----PSLTLSPSSE--GSCSLKNNFLSLGGGACATCPLSTNILLGTSVA 59

Qy 108 PQrtqGVRGAIWVRSSSIPLRAASPTPWGTGTYSSAPPLALPG 148  
Db 60 TFSAMIMGRGLWTSPSYIVPRRGPSSE--TRLGAPVPTGSPG 98

**RESULT 6**

T19006 ankyrin related protein C06C3.1 - Caenorhabditis elegans  
N;Contains: myosin-light-chain-phosphatase (EC 3.1.3.53)  
C;Species: Caenorhabditis elegans  
C;Accession: T15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19006; T22086  
R;Burkitt, M.  
R;Matthews, P.  
submitted to the EMBL Data Library, August 1994  
A;Reference number: Z19058  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1016 <WIL>  
A;Cross-references: UNIPROT:Q17718; EMBL:Z36719; PIDN:CAA85318.1; GSPPDB:GN00020; CESPP:CO  
A;Experimental source: clone C06C3  
A;Accession: T19006  
A;Accession: T22086  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1016 <W12>  
A;Cross-references: EMBL:Z47809; PIDN:CAA87782.1; GSPPDB:GN00020; CESPP:C06C3.1  
A;Experimental source: clone F42A8  
C;Genetics:  
A;Map position: 2  
A;Introns: 27/3; 94/3; 279/3; 352/2; 462/2; 523/3; 569/2; 657/2; 718/3; 766/3; 833/3; 89  
C;Keywords: phosphoric monoester hydrolase

Query Match 10.5%; Score 90; DB 2; Length 1016;  
Best Local Similarity 22.3%; Pred. No. 5.3;  
Matches 39; Conservative 22; Mismatches 60; Indels 54; Gaps 6;

Qy 6 LKKKTPVTPVTTSSSRVAV-RSLVVEFTCCRAGLDWTCARRRLPSGRNLFEVDVLSI 64  
Db 439 LKNGKEISPLRSETTSSRSTSITSLSDGVT-----DRSSSGRETSAEAMS-- 482

Qy 65 PRLVGPRAGPLSPGTLG-----PSNAMRAGG-----RD 94  
Db 483 -----EPASSASTGTTSSSSRFTSSTPSQRSGSVHETPRASMENSSVGQDNV 536

Qy 95 GSCLUPVALGLAGAQPQPGYGRAIWRSISPLRAASPTSMGTYRSSAAPLEALPGP 149  
Db 537 SATIPV-VPLSAPPKAHVHQSPSSWNRGVPLOSSRTSSVTRSSVTPVSEIISPP 590

**RESULT 7**

2

Page 4

A; Cross-references: GDB:1202555 ; OMIM:164176			
A; Map Position: 19pter-19pter			
A; Introns: 221/3; 269/1; 318/3			
C; Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology			
C; Keywords: alternative splicing; B-cell; DNA binding; homeobox; nucleus; transcription			
P; 298-354/Domain: POU domain homology <Hox>			
Query Match Score 10.4%; Score 89; DB 1; Length 479;			
Best Local Similarity 23.9%; Pred. No. 3;			
Matches 32; Conservative 10; Mismatches 28; Indels 64; Gaps 6;			
Qy 43 WYCARRIER-----LPS-GRNLIEDVDSLSPRLVCPRAFGCLSP----- 78			
Db 344 WFCNREQRQEKKRINPCSAAPMLPSGR-----PASYSPHMTPQGEGGTPLSQASSSLST 399			
Qy 79 -----GTGJPSMAMRAAGRDGCGLPV----- 104			
Db 400 VTTLSSAVGILHPSRTAGGGGGGAPPLNSIPSVTPPPATNSTNFSPOGSHSAIGL 459			
Qy 105 AGAAPOTPGVGRAIW 118			
Db 460 SGL--NPSTGPGLW 471			
RESULT 8			
F75311 ABC transporter, ATP-binding protein - Deinococcus radiodurans (strain R1)			
C;Species: Deinococcus radiodurans			
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004			
C;Accession: F75311			
C;R.White, O.; Bisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, R.J.; Dodson, R.J.; Zalewski, C.; Ma, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Venet, J.C.; Fraser, C.M.			
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.			
Science 286, 1571-1577, 1999			
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.			
A;Accession: F75311			
A;Status: preliminary			
A;Cross-references: 1-552 <WHI>			
A;Cross-references: UNIPROT:Q9RS9H ; GB:AE002048; GB:AE000513; NID:6459929; PIDN:AAFI168			
A;Experimental source: strain R1			
C;Genetics:			
A;Gene: DR2145			
A;Map position: 1			
Query Match Score 10.4%; Score 89; DB 2; Length 552;			
Best Local Similarity 29.5%; Pred. No. 3.5;			
Matches 49; Conservative 10; Mismatches 55; Indels 52; Gaps 9;			
Qy 18 HRTSSSRVAVRSLVEFTCCRAGALDWCVARRLPGRNILEVDVLSSPRLVGPAGPG---- 75			
Db 119 HAAAARRA-----SCARERHPD-VCAR----TGSAAFRGAGAYPE---RGGAAGR 162			
Qy 76 --LSPOTLGPSMAMRAAGGDGSCLPVALLAGAQTPGVGRATW-----VRSSIP 124			
Db 163 GVLHSRLGTDRAARRAGGRR-TPAAVG-SDAARTTRPVGGRDYDTRRGALAGARLSPH 220			
Qy 125 LRAASPTSWGTYRSSA-----PILLEALPC 149			
Db 221 LARAPPDGAPRDRARAAPAAARRIHRWPERSGPRPRRAAAGP 266			
RESULT 9			
S22544 transcription factor Oct-2, splice form Oct-2.5 - mouse			
N;Alternative names: NF- $\alpha$ ; Oct-2; transcription factor Oct-2b			
C;Species: Mus musculus (house mouse)			
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004			
C;Accession: S22544 ; A60085 ; S09240			
R;Nirth, T.; Priess, A.; Anweiler, A.; Zwilling, S.; Oeler, B.			
Nucleic Acids Res. 19, 43-51, 1991			

C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation; F;262-263.Domain: POU domain homology <POU> F;298-354.Domain: homeobox homology <HOX>	Query Match	10.2%; Score 87.5; DB 1; Length 478; Best Local Similarity 26.7%; Pred. No. 4.1.; Mismatches 9; Indels 75; Gaps 10; Matches 44; Conservative 9;
Db	43 WVCARRER-----LPS-GRNLEVDVSLSPRLVPGPAGPGLSP-----78 344 WFCNRQKEKRINPCSAAPMLPSGK----PASYSPHLVTQGGAGTPLSQASSLSNT 399	
Qy	79 -----GTIGPSMAMRAAGGRDGSCLPVALGLAGAQTPGVGRAIVRSSPLRLAASP 130	
Db	400 VTLSSSAVGTLHPS---RTAGG----GAAGGGAAPPL-----NSTP--SVTP 437	
Qy	131 TSWGTYRSSLPL-----LEALPGPWRMASGFWKATMQ 163	
Db	438 PPPATTNSTNFNSPQGSHSAIGLGLNESTGP----GLWNPAPYQ 477	
<b>RESULT 11</b>		
D75622	hypothetical protein - Deinococcus radiodurans (strain R1)	
C;Species: Deinococcus radiodurans		
C;Accession: D75622	#sequence_revision 03-Dec-1999 #text_change 09-Jul-2004	
R;White, O.; Eisen, J.A.; Heidelberg, J.B.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Ma, S.; Smith, H.O.; Ventre, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, S.; Smith, H.O.; Ventre, J.C.; Fraser, C.M.	Science 286, 1571-1577, 1999.	
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.		
A;Reference number: A75250; PMID:20036896; PMID:10567266		
A;Accession: D75622		
A;Status: Preliminary		
A;Molecule type: DNA		
A;Residue: 1-156 <WHI>		
A;Cross-references: UNIPROT:Q9RZ55; GB:AE001826; NID:g6460827; PID:AAF12630.1; PID:g6460827		
A;Experimental source: strain R1		
C;Genetics:		
A;Gene: DRB0039		
A;Map position: megaplasmid		
A;Genome: Plasmid		
A;Note: Plasmid MP1		
<b>RESULT 12</b>		
B86483	protein F5J5_15 [imported] - Arabidopsis thaliana	
C;Species: Arabidopsis thaliana (mouse-ear cress)		
C;Accession: B86483	#sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	
R;Theologis, A.; Baker, J.R.; Palm, C.J.; Fedderspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Hansen, N.P.; Hughes, B.; Huizar, L.	Nature 408, 816-820, 2000.	
A;Author(s): Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.J.; Kim, C.		
Qy	10 TNVPTVVAHTSSRVAVRSVLVEFTCCRAGALDWVCCRERLPSGRNLEVDVSLSPRLVVG 69	
Db	286 SNATDLIRRVTQPLGRLRK----TVSKVGAEDW----APGSATVSGDELEYTAVTNPYAQ 339	
Qy	70 PRAG-PGLSPOTLGPMAMRAGGRDOSCLPVALGLGAPQT----PGVGRALWVR 120	
Db	340 PLAGVQLDPLPAGTEFVSADGG----ALLGAASAQAVWTLGDLPAGATRTLRL 392	
Qy	121 -----SSPL----RAASPTSWGTYRSSLPL-----EALPG 148	
Db	393 VRVGDVRDDOBLRNVFEITSSPLPFLHSNDASAVWT----APLISKTLDRDAPG 448	

RESULT 14

A39481 serum response factor-related protein 2 - human  
N:Alternate names: myocyte-specific enhancer factor xnef2; RSRF2  
C:Species: Homo sapiens (man)  
C:Date: 03-Aug-1992 #sequence revision 03-Aug-1992 #text\_change 09-Jul-2004  
C:Accession: A39481; S25832; S24468  
R;Pollock, R.; Treisman, R.  
Genes Dev. 5, 2327-2341, 1991  
A;Title: Human SRF-related proteins: DNA-binding properties and potential regulatory targets  
A;Reference number: A39481; MUID:92084105; PMID:1748287  
A;Accession: A39481  
A;Molecule type: mRNA  
A;Residues: 1-365 <YUY>  
A;Cross-references: UNIPROT:002080; EMBL:X68502; NID:G37991; PIDN:CAA48515.1; PID:937992  
C;Keywords: DNA binding; homodimer; transcription factor  
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match, Score 9.9%; Best Local Similarity 9.9%; Pred. No. 5.1; Length 365;  
Matches 36; Conservative 10; Mismatches 47; Indels 50; Gaps 7;

Qy 47 RRBDPSGRNLLEVDSLSPVLPGVPPGPGUSPGLGSPAMRAAGGRDSCLPVALG--L 104  
Db 207 RRSIDLPGG-----LAGFRGG-----INTSSRLS---GLQNPCSTATPGPL 245

Qy 105 AGAPOTPG---VGRAIWWR-----SSIPYRASPTNSGYTSSAAPPLEA 145  
Db 246 GSFPPLPGPPVGABAARRVPQAPPRPPQASSSLASLRPPGAPTFLRSPPIPOS 305

Qy 146 LPGPKRMAASGF-----WKT A 160  
Db 306 SPGPHQSLCGLGPPCAAGCPMPATA 328

---

RESULT 15

EDBEXD

1immediate-early protein RL2 - human herpesvirus 2 (strain HG52)

N:Alternate names: RL2 protein  
C:Species: human herpesvirus 2  
A;Note: host Homo sapiens (man)  
C:Accession: JQ1501  
R;McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.  
J. Gen. Virol. 72, 3057-3075, 1991  
A;Title: Comparative sequence analysis of the long repeat regions and adjoining parts of  
A;Reference number: JQ1194; MUID:9211549; PMID:166297  
A;Accession: JQ1501  
A;Molecule type: DNA  
A;Residues: 1-825 <MCG>  
A;Cross-references: UNIPROT:P28284; GB:D10471; DDBJ:D01128; NID:g221784; PIDN:BAA23427.1  
C;Genetics:  
A;Gene: RL2  
A;Introns: 25/3; 22/1  
C;SuperFamily: herpesvirus immediate-early protein IRI10; RING finger homology  
C;Keywords: DNA binding; immediate-early protein; tandem repeat; transcription regulation  
F;122-172/Region: RING finger homology <RNG>  
F;226-316/Region: zinc finger C3HC4 motif  
F;589-623/Region: 5-residue repeats (A-S-S-S-S)

Query Match, Score 9.9%; Best Local Similarity 27.0%; Pred. No. 12; Length 825;  
Matches 40; Conservative 11; Mismatches 63; Indels 34; Gaps 5;

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw mode!

Run on: August 25, 2005, 03:18:18 ; Search time 171 Seconds  
 (without alignments/sec 491.117 Million cell updates/sec)

Title: US-10-664-038-1

Perfect score: 857

Sequence: 1 ARILNLKKCTNVTPTVAHRT.....ALPGPWRMASGFWKTTATMQQ 164

Scoring table: BIOSUM62

dapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:  
 1: uniprot\_sprot:/\*  
 2: uniprot\_trembl:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description	
1	577	67.3	125	2	Q66VR7	Q66vr7 hepatitis C	Q66vr7 hepatitis C	
2	557	65.0	125	2	Q66VN9	Q66vn9 hepatitis C	Q66vn9 hepatitis C	
3	556	64.9	125	2	Q66VKL	Q66vkl hepatitis C	Q66vkl hepatitis C	
4	432	50.5	126	2	Q66VN1	Q66vn1 hepatitis C	Q66vn1 hepatitis C	
5	403	47.0	122	2	Q00687	Q00687 hepatitis C	Q00687 hepatitis C	
6	374	43.6	95	2	Q68874	Q68874 hepatitis C	Q68874 hepatitis C	
7	350	40.9	134	2	Q70GD2	Q70gd2 hepatitis C	Q70gd2 hepatitis C	
8	346	40.4	102	2	Q68361	Q68361 hepatitis C	Q68361 hepatitis C	
9	336	39.2	102	2	Q68358	Q68358 hepatitis C	Q68358 hepatitis C	
10	331	37.7	102	2	Q68356	Q68356 hepatitis C	Q68356 hepatitis C	
11	319	37.2	102	2	Q68359	Q68359 hepatitis C	Q68359 hepatitis C	
12	305	35.6	100	2	Q68360	Q68360 hepatitis C	Q68360 hepatitis C	
13	301	35.1	71	2	Q91AW2	Q91aw2 hepatitis C	Q91aw2 hepatitis C	
14	296	34.5	71	2	Q91AW1	Q91aw1 hepatitis C	Q91aw1 hepatitis C	
15	287	33.5	102	2	Q68365	Q68365 hepatitis C	Q68365 hepatitis C	
16	285	33.3	102	2	Q68366	Q68366 hepatitis C	Q68366 hepatitis C	
17	262	30.6	102	2	Q68363	Q68363 hepatitis C	Q68363 hepatitis C	
18	262	30.6	102	2	Q68368	Q68368 hepatitis C	Q68368 hepatitis C	
19	254	29.6	102	2	Q68367	Q68367 hepatitis C	Q68367 hepatitis C	
20	225	26.3	80	2	Q81289	Q81289 hepatitis C	Q81289 hepatitis C	
21	220	25.7	119	2	Q86686	Q86686 hepatitis C	Q86686 hepatitis C	
22	148	17.3	53	2	Q91RN9	Q91rn9 hepatitis C	Q91rn9 hepatitis C	
23	142	5	16.6	125	2	Q66VTO	Q66vto hepatitis C	Q66vto hepatitis C
24	142	5	16.6	125	2	Q66VY1	Q66vy1 hepatitis C	Q66vy1 hepatitis C
25	142	5	16.6	125	2	Q66VY3	Q66vy3 hepatitis C	Q66vy3 hepatitis C
26	137	5	16.0	125	2	Q66VP7	Q66vp7 hepatitis C	Q66vp7 hepatitis C
27	137	5	16.0	125	2	Q66VQS	Q66vqs hepatitis C	Q66vqs hepatitis C
28	123	5	14.4	210	2	P6NYL1	Oenyl1 homo sapien	P87760 hepatitis C
29	109	5	12.8	108	2	P87760	P87760 hepatitis C	P87761 hepatitis C
30	109	5	12.8	108	2	P87761	P87761 hepatitis C	Q16992 anthopleura
31	102	5	12.0	1	LWA_ANTEL			

## ALIGNMENTS

RESULT 1											
Q66VR7		PRELIMINARY;		PRT:		125 AA.					
ID Q66VR7,				AC Q66VR7,				DT 25-OCT-2004 (TREMBLrel. 28, Created)			
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)				DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)				DB Core protein (fragment).			
DB Hepatitis C virus.				RA Hepatitis C virus.				RA Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;			
OC Hepacivirus.				RA Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;				OC Hepacivirus.			
OX NCBI_TaxID=11103;				OX NCBI_TaxID=11103;				RN [1] _			
RP SEQUENCE FROM N.A.				RP PubMedID=15302945; DOI=10.1099/vir.0.80071-0;				RX PubRef=S.; Blashev A.; Perret M.; Sodoyer M.; Pouzol S.; Scoazec J.Y.; Bonnaud B.; Diaz O.; Paranhos-Baccala G.; Lotteau V.; Andre P.;			
RT "Expression of hepatitis C virus proteins in epithelial intestinal cells in vivo.";				RT RL J. Gen. Virol. 85:2515-2523 (2004).				RT DR EMBL; AY600651; AAU11610_1; -.			
FT NONTER 125 125				FT NONTER 125 125 MW: 3B44B44EF7601358 CRC64;				SEQUENCE 125 AA; 12976 MW;			
SQ				Query Match 67.3%; Score 577; DB 2; Length 125;				Best Local Similarity 91.1%; Pred. No. 3.5e-38;			
Matches 113; Conservative 1; Mismatches 10; Indels 0; Gaps 0;				Matches 113; Conservative 1; Mismatches 10; Indels 0; Gaps 0;				SEQUENCE 125 AA; 12976 MW;			
Qy 1 ARILNPKKTKNTVTPYTAHTTSSSRVAVRSLVYFTCCAGALDWCRARRLFPSGRNIEVD 60				Qy 1 ARILNPKKTKNTVTPYTAHTTSSSRVAVRSLVYFTCCAGALDWCRARRLFPSGRNIEVD 60				Db 2 ARILNPKKTPNTPYTAHTTSSSRVAVRSLVYFTCCAGALDWCRARRLFPSGRNIEVD 61			
Qy 61 VSLSPRLVGPRAKGPSGTLGPSPGTLGPSPMSMAMRAAGGRDGSCLPVALGLAGAQPOTPGVGRAIWVR 120				Qy 61 VSLSPRLVGPRAKGPSGTLGPSPGTLGPSPMSMAMRAAGGRDGSCLPVALGLAGAQPOTPGVGRAIWVR 120				Db 62 VSLSPRLVGPRAKGPSGTLGPSPGTLGPSPMSMAMRAAGGRDGSCLPVALGLAGAQPOTPGVGRAIWVR 121			
Qy 121 SSIP 124				Qy 121 SSIP 124				Db 122 SSIP 125			
Q66VM9 PRELIMINARY;				Q66VM9 PRELIMINARY;				PRT: 125 AA.			
AC Q66VM9;				AC Q66VM9;				DT 25-OCT-2004 (TREMBLrel. 28, Created)			
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)				DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)				DB Core protein (fragment).			
DE Hepacivirus.				DE Hepacivirus.				OS Hepatitis C virus.			
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;				OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;				OC Hepacivirus.			
OC Hepacivirus.				OC Hepacivirus.				OC Hepacivirus.			
OX NCBI_TaxID=11103;				OX NCBI_TaxID=11103;				OX NCBI_TaxID=11103;			



PT	NON_TER	122	122 AA;	122 MW;	PB5C7A7F76AE29AC CRC64;		OX	NCBI_TaxID=41856;
SQ	SEQUENCE	122 AA;					RN	SEQUENCE FROM N.A.
	Query Match	47 0%	Score 403;	DB 2;	Length 122;		RX	PubMed:14718617; DOI=10.1099/vir.0.19472-0;
	Best Local Similarity	68.6%	Pred. No. 1.9e-24;				RA	Colina R., Casane D., Vasquez S., Garcia L., Chunga A., Romero H., Kahn B., Cristina J.,
	Matches	83;	Conservative	9;	Mismatches 29;	Indels 0;	RA	"Genetic analysis of Hepatitis C viruses in Peruvian patients.";
Qy	6	LKKTKNVTPYTAHRTSSSRVAVRSLVEFTCRAGALDWCARRELPSGRNLVDVSLSGP 65					RA	J. Gen. Virol. 85:31-37 (2004).
Db	1	LKEKEPKETPSVAKHTSSSRVADRSLVEYTCCRAGAHDWCARRELPSGRNLVDVSLSGP 60					DR	EMBL: AJ582228; CABE46584.1; -
Qy	66	RLVGRAGRGPSPGTPGSPMAMRAGGRDSCSCLPVALGLAGAPTPGYGRAINWRSIPL 125					DR	GO: 0019028; C:viral capsid; IEA,
Db	61	RRVGAKAGPGLUSPGRGPSPTRGPSPMVTRAGGGGSCPHAPVHLGAQMTPGEPAIWVKSSIFS 120					DR	DR InterPro: IPR02522; HCV capsid.
Qy	126	R 126					DR	PFam: PF01543; HCV_capsid; 1.
Db	121	R 121					FT	NON_TER 134 AA;
SQ	SEQUENCE	134 AA;					FT	14214 MW;
							SQ	85D9075FB32CABAC CRC64;
							Query Match	40.9%; Score 350.5;
							Best Local Similarity	85.4%; Pred. No. 2.9e-20;
							Matches	70; Conservative 1; Mismatches 6; Indels 5; Gaps 1;
Qy	52	PSGRNLVDVSLSPLRVGSPRAGPGLSPCTLGPSMAAAGGRDGSCLPVALGLAGAPTP 111					Qy	52 PSGRNLVDVSLSPLRVGSPRAGPGLSPCTLGPSMAAAGGRDGSCLPVALGLAGAPTP 111
ID	Q68874	PRELIMINARY;					Db	58 PROB: ---QTPPRHYGSPRACPGLSFGTGLGPSMAAAGGRDGSCLPVALGLAGAPTP 112
AC	Q68874;	PRT;	95 AA.				Qy	112 GYGRAINWRSIPLRAASPTSW 133
DT	01-NOV-1996	(TREMBLrel. 01, Created)					Db	113 GYGRAINWRSIPLRAASPTSW 134
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)						
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)						
DE	Core protein (Fragment).						DE	Core protein (Fragment).
OS	Hepatitis C virus.						OS	Hepatitis C virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;						OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
NCBI_TaxID	11103;						OC	Hepacivirus.
RN	[1]						OC	Hepacivirus.
RP	SEQUENCE FROM N.A.						RN	[1]
RX	MEDLINE:95270681; PubMed:7751366;						RP	SEQUENCE FROM N.A.
RA	Peucht H.H., Zoellner B., Poliakwa S., Laufs R.; Laufs R.;						RC	STRAIN=NEU03;
RT	"Study on reliability of commercially available hepatitis C virus antibody tests."						RC	Medline:96030859; PubMed:7595353;
RT	antibody tests.";						RC	RA Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RL	J. Clin. Microbiol. 33:620-624 (1995).						RC	The Int'l. HCV Collaborative Study Group;
RN	[2]						RC	RT Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification.";
RA	Peucht H.;						RL	RT J. Gen. Virol. 76:2433-2507 (1995).
RL	Submitted (APR-1994) to the EMBL/GenBank/DBDJ databases.						RL	RA Mellor J., U31237; MAA75001.1; -.
PIR	X88950; CAAS5547.1; -.						DR	RA FT NON_TER 1 1
DR	S4213; S4423.						FT	FT NON_TER 102 102
PT	NON_TER 1						SQ	SQ SEQUENCE 102 AA; 10446 MW; D3C18355326B429D CRC64;
SQ	95 AA;	9702 MW;	647C80587C6F892F CRC64;				Query Match	40.4%; Score 346;
							Best Local Similarity	69.9%; Pred. No. 5.1e-20;
							Matches	72; Conservative 6; Mismatches 23; Indels 2; Gaps 2;
Qy	11	NVTPTVAHTSSSVAVRSLVEFTCRAGALDWCARRELPSGRNLVDVSLSGP 70					Qy	9 KTNVTPYVAHTSSSVAVRSLVEFTCRAGALDWCARRELPSGRNLVDVSLSGP 68
Db	2	NVTPTAAHTSSSVAVRSLVEFTCRAGALDWCARRELPSGRNLVDVSLSGP 61					Db	1 KTNVTPYAAHTSSSVAVRSLVEFTCRAGALDWCARRELPSGRNLVDVSLSGP 60
Qy	71	RAGPGLSPCTLGPSMAMRAGGRGSCUPLVALGL 104					Qy	69 GPRAGPGLSPCTLGPSMAMRAGGRGSCUPLVALGL 110
Db	62	RAGPGLSPCTLGPSMAMRAGGRGSCUPLVALGL 95					Db	61 SPRAGPGLSPCTLGPSMAMRAGGRGSCUPLVALGL 102
RESULT	7							
O7GD2	PRELIMINARY;							
AC	O7GD2;	PRT;	134 AA.					
DT	05-JUL-2004	(TREMBLrel. 27, Created)						
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)						
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)						
DE	Core protein (Fragment).							
OS	Hepatitis C virus type 1.							
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;							
OC	Hepacivirus.							

DT	01-NOV-1996 (TREMBLrel. 01, Created)	Db	61 GAKGGPGLSPGTIGPSMVTRAAGGQQGSCPAAAPVQLG 98
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DE	Core protein (Fragment).		
OS	Hepatitis C virus.		
Virus;	ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepacivirus.		
NCBI_TaxID=11103;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SUBAH345;		
RX	MEDLINE=96030859; PubMed=7595353;		
RA	Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,		
RA	The Int'l HCV Collaborative Study Group;		
RT	"Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification."		
RT	J. Gen. Virol. 76:2493-2507 (1995).		
DR	EMBL: U31254; AAA75058.1; -.		
FT	NON_TER 1 1		
FT	NON_TER 102 102 AA: 10407 MW; E3B0611PF5880B35 CRC64;		
SQ	SEQUENCE 102 AA: 102 AA; 10407 MW; E3B0611PF5880B35 CRC64;		
Query Match	39.2%; Score 336; DB 2; Length 102;		
Best Local Similarity	69.7%; Pred. No. 3.1e-19;		
Matches	69; Conservative 2; Mismatches 28; Indels 0; Gaps 0;		
Qy	9 KTNVPTVAHRTSSSRVAVSLPEFCRAGALDNYCARERLPGRNLDEVDSLSPLRV 68		
Db	1 KPNVPTTAAPWTSRVAVSLPEFCRAGALDNYCARELSPRLSGRNLVEGVNLSPSA 60		
Qy	69 GPRAGPGLSPGTIGPSMAMRAAGRDGSCLPVLAGA 107		
Db	61 GPRAGPGLSPGTIGPSMAMRAAGRDGSCLPVLAGA 106		
RESULT 10			
ID	Q68356 PRELIMINARY;	PRT;	102 AA.
AC	068356;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DB	Core protein (Fragment).		
OS	Hepatitis C virus.		
Virus;	ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepacivirus.		
NCBI_TaxID=11103;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EU0K43;		
RX	MEDLINE=96030859; PubMed=7595353;		
RA	Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,		
RA	The Int'l HCV Collaborative Study Group;		
RT	"Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification."		
RT	J. Gen. Virol. 76:2493-2507 (1995).		
DR	EMBL: U31252; AAA75056.1; -.		
FT	NON_TER 1 1		
FT	NON_TER 102 102 AA: 10539 MW; 30BBC3CDB0F80E96 CRC64;		
SQ	SEQUENCE 102 AA: 10539 MW; 30BBC3CDB0F80E96 CRC64;		
Query Match	38.7%; Score 331.5; DB 2; Length 102;		
Best Local Similarity	68.4%; Pred. No 7.1e-19;		
Matches	67; Conservative 7; Mismatches 21; Indels 3; Gaps 1;		
Qy	9 KTNVPTVAHRTSSSRVAVSLPEFCRAGALDNYCARERLPGRNLDEVDSLSPLRV 68		
Db	1 KPKETPSVAVRTSSSRVAVSLPEFCRAGALDNYCARELSPRLSGRNLDEVDSLSPLRV 60		
Qy	69 GPRAGPGLSPGTIGPSMAMRAAGRDGSCLPVLAG 103		
Db	61 GPRAGPGLSPGTIGPSMAMRAAGRDGSCLPVLAG 103		
RESULT 11			
ID	Q68356 PRELIMINARY;	PRT;	100 AA.
AC	068356;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DB	Core protein (Fragment).		
OS	Hepatitis C virus.		
Virus;	ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepacivirus.		
NCBI_TaxID=11103;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EU0G28;		
RX	MEDLINE=96030859; PubMed=7595353;		
RA	Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,		
RA	The Int'l HCV Collaborative Study Group;		
RT	"Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification."		
RT	J. Gen. Virol. 76:2493-2507 (1995).		
DR	EMBL: U31256; AAA75060.1; -.		
FT	NON_TER 1 1		
FT	NON_TER 100 100 AA: 10414 MW; 727D8920E9BE281C CRC64;		
SQ	SEQUENCE 100 AA: 10414 MW; 727D8920E9BE281C CRC64;		
Query Match	35.6%; Score 305; DB 2; Length 100;		
Best Local Similarity	67.0%; Pred. No. 8.6e-17;		
Matches	65; Conservative 2; Mismatches 30; Indels 0; Gaps 0;		

Qy 11 NVTPTVAAHRTSSSRVAVRSVLVEFTCCRAGALDWCCRRLPGRNLEVDSLSPLRVGP 70  
 Db 1 NXIPTAARTLSSSRVAAARSLAEFTCCRAGAPGWCARLGLRSLRNVLVEGANLPFRRAAP 60

Qy 71 RADPGISPGTIGPSMAMRAAGGRDGSCLPVALGLAGA 107  
 Db 61 RADGXSLGIGLGPFWAMRAAGGGCSCHPRPLGLGA 97

RESULT 13  
 Q9IAW2 :PRELIMINARY; PRT; 71 AA.  
 ID Q9IAW2 ;  
 AC 09IAW2 ;  
 DT 01-DBC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DB Polyprotein (Fragment).  
 OS Hepatitis C virus.  
 OC ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP - SEQUENCE FROM N.A. MEDLINE=2194401; PubMed=11926855;  
 RX RA - Ross R.S., Viazov S., Thormahlen M., Bartz L., Tamm J., Rautenberg P.,  
 RA , Roggendorf M., Deister A.;  
 RT "Risk of hepatitis C virus transmission from an infected gynecologist  
 to patients: results of a 7-year retrospective investigation.";  
 RL Arch. Intern. Med. 162:805-810(2002).  
 [2]  
 SEQUENCE FROM N.A.  
 RA Ross S., Viazov S., Roggendorf M.;  
 DR Submitted (PBB-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; AP250919; AAL1060.1; -  
 KW Polyprotein.  
 PT NON\_TER 1 1  
 SQ SEQUENCE 71 AA; 7096 MW; E7E89BB9D1AF95AC CRC64;  
 Query Match 35.1%; Score 301; DB 2; Length 71;  
 Best Local Similarity 88.1%; Pred. No. 1..3e-16;  
 Matches 59; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 FT NON\_TER 71 1  
 SQ 44 VCARRERLSSGRNLLEVDSLSPLRVGPAGLSPGTLGPSMAMRAAGGRDGSCLPVALG 103  
 5 VCAARRGRLLSDRNLEVDVSLSPRHYGPRAAGPGLSPGTLGPSMAMRAVGQDGSCPPAALG 64

Qy 104 LAGAPQT 110  
 Db 65 Lvdapqt 71

RESULT 14  
 Q9IAW1 :PRELIMINARY; PRT; 71 AA.  
 ID Q9IAW1 ;  
 AC 09IAW1 ;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DB Polyprotein (Fragment).  
 OS Hepatitis C virus.  
 OC ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP - SEQUENCE FROM N.A. MEDLINE=2194401; PubMed=11926855;  
 RX RA - Ross R.S., Viazov S., Thormahlen M., Bartz L., Tamm J., Rautenberg P.,  
 RA , Roggendorf M., Deister A.;  
 RT "Risk of hepatitis C virus transmission from an infected gynecologist  
 to patients: results of a 7-year retrospective investigation.";  
 RL Arch. Intern. Med. 162:805-810(2002).

[2]  
 SEQUENCE FROM N.A. Ross S., Viazov S., Roggendorf M.;  
 DR Submitted (PBB-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; AR350920; AAL11061.1; -  
 KW Polyprotein.  
 PT NON\_TER 1 1  
 SQ SEQUENCE 71 AA; 7126 MW; 918954D735E26CA5 CRC64;  
 Query Match 34.5%; Score 296; DB 2; Length 71;  
 Best Local Similarity 88.1%; Pred. No. 3..2e-16;  
 Matches 59; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 FT NON\_TER 71 1  
 SQ 44 VCARRERLSSGRNLLEVDSLSPLRVGPAGLSPGTLGPSMAMRAAGGRDGSCLPVALG 103  
 5 VCAARRGRLLSDRNLEVDVSLSPRHYGPRAAGPGLSPGTLGPSMAMRAVGQDGSCPPAALG 64

Qy 104 LAGAPQT 110  
 Db 65 Lvdapqt 71

RESULT 15  
 Q9IAW2 :PRELIMINARY; PRT; 102 AA.  
 ID Q9IAW2 ;  
 AC 09IAW2 ;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Core protein (Fragment).  
 OS Hepatitis C virus.  
 OC ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP - SEQUENCE FROM N.A. MEDLINE=96030859; PubMed=7595353;  
 RX RA - Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,  
 RA The Int'l HCV Collaborative Study Group;  
 RT "Investigation of the pattern of hepatitis C virus sequence diversity  
 in different geographical regions: implications for virus  
 classification.",  
 RL J. Gen. Virol. 76:2493-2507(1995).  
 DR EMBL; U31261; AAA75005.1; -  
 PT NON\_TER 1 1  
 SQ SEQUENCE 102 AA; 10615 MW; 045EBE22E77C8AA00 CRC64;  
 Query Match 33.5%; Score 287; DB 2; Length 102;  
 Best Local Similarity 59.8%; Pred. No. 2..3e-15;  
 Matches 61; Conservative 6; Mismatches 35; Indels 0; Gaps 0;

Qy 9 KTNVTPVAAHRTSSSRVAVRSVLVEFTCCRAGALDWCCRRLPGRNLEVDSLSPLRV 68  
 Db 1 KPKETPAAHRTSSRAVRSSVEYTCCRAGPWTQQQLQNPAKEXGANQYPRRA 60

Qy 69 GPRAGPGLSPGTIGPSMAMRAAGGRDGSCLPVALGLAGAPQT 110  
 Db 61 TRRAAPGNSLGTNGMRAAGQDSCPPVALAHTGAPMT 102

Search completed: August 25, 2005, 03:36:22  
 Job time : 173 sec

Result No.	Score	Query	Match	Length	DB ID	Description
1	857	100.0	164	4	US-09-644-987-1	Sequence 1, Appli
2	324	37.8	115	4	US-09-878-281A-148	Sequence 148, APP
3	100	11.7	236	4	US-09-252-991A-27018	Sequence 27618, A
4	99.5	11.6	263	4	US-09-252-991A-16306	Sequence 16906, A
5	98	11.4	162	4	US-09-252-991A-18207	Sequence 18207, A
6	98	11.4	202	4	US-09-252-991A-24790	Sequence 24790, A
7	97	11.3	231	4	US-09-252-991A-22801	Sequence 22801, A
8	96	11.2	709	4	US-09-252-991A-27305	Sequence 27305, A
9	95.5	11.1	328	4	US-09-232-991A-17729	Sequence 17729, A
10	95	11.1	371	4	US-09-252-991A-25006	Sequence 25006, A
11	93	10.9	213	4	US-09-252-991A-24558	Sequence 24558, A
12	92.5	10.8	254	4	US-09-232-991A-23111	Sequence 23111, A
13	91.5	10.7	461	4	US-09-232-991A-24717	Sequence 24717, A
14	91	10.6	253	4	US-09-252-991A-29632	Sequence 29632, A
15	91	10.6	1027	4	US-09-232-991A-26216	Sequence 26216, A
16	90	10.5	220	4	US-09-232-991A-16739	Sequence 16739, A
17	90	10.5	242	4	US-09-252-991A-31425	Sequence 31425, A
18	89.5	10.4	138	4	US-09-232-991A-26931	Sequence 26931, A
19	89.5	10.4	1209	4	US-09-252-991A-25844	Sequence 25844, A
20	89.5	10.4	1225	4	US-09-252-991A-25018	Sequence 25018, A
21	89	10.4	348	4	US-09-232-991A-25605	Sequence 25605, A
22	89	10.4	425	4	US-09-232-991A-20467	Sequence 20467, A
23	89	10.4	470	4	US-09-949-016-9859	Sequence 9859, AP
24	88.5	10.3	171	4	US-09-232-991A-23116	Sequence 23216, A
25	88	10.3	146	4	US-09-232-991A-28752	Sequence 27352, A
26	88	10.3	504	4	US-09-232-991A-28442	Sequence 28424, A
27	87.5	10.2	205	4	US-09-252-991A-17563	Sequence 17563, A

## ALIGNMENTS

RESULT 1  
US-09-644-987-1  
; Sequence 1, Application US/09644987  
; Patent No. 6803214  
; GENERAL INFORMATION:  
; APPLICANT: MAVRONARA, PENELOPE  
; VARALIOTI, AGORITSA  
; GEORGOPOLOU, URANIA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND NEW POLYPEPTIDES ASSOCIATED WITH HEPATITIS C VIRUS CORE GENE  
; TITLE OF INVENTION: AND/OVERLAPPING WITH HEPATITIS C VIRUS CORE GENE  
; TITLE OF INVENTION: PRODUCTS  
; FILE REFERENCE: 03495-0194-00000  
; CURRENT APPLICATION NUMBER: US-09-644, 987  
; CURRENT FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/151, 074  
; PRIOR FILING DATE: 1999-08-27  
; NUMBER OF SEQ ID NOS: 16  
; SEQ ID NO: 1  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
; US-09-644-987-1  
Query Match 100.0%; Score 857; DB 4;  
Best Local Similarity 100.0%; Pred. No. 5.8e-82;  
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CURRENT APPLICATION NUMBER: US/09/878,281A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 148  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-09-878-281A-148

Query Match 37.8%; Score 324; DB 4; Length 115;  
Best Local Similarity 63.0%; Pred. No. 2.e-26;  
Matches 68; Conservative 4; Mismatches 34; Indels 2; Gaps 1;

Qy 7 KCTTNVTPTVAHRTSSVAVRSSEVEFTCRAGLDWVCCRERLPSGRNLLEVDSLSPR 66  
Db 10 KTRNTNTP-GHRTLSSQANVRLSVELFTCYHAGAPSIVCQCARLPSGRNLAVGANPSPG 67

Qy 67 LVGPRAGPGLSPGTLPSPMAMRAGGRDGSCLPVALGLAGAPQTPGVC 114  
Db 68 RAERAGPGLSPGTLPSPMAMRAGGQGSCPAAALARGAQNTPGAG 115

---

RESULT 3  
US-09-252-991A-27618  
; Sequence 27618; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.13  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27618  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27618

Query Match 11.7%; Score 100; DB 4; Length 236;  
Best Local Similarity 28.3%; Pred. No. 0.015; Mismatches 13; Indels 52; Gaps 8;

Qy 34 TCCAGALDWVCCRERLPSGRNLLEVDSLSPRVLGPGRGSPGTLPSPMAMRAAGGR 93  
Db 46 TCCRSPGPAGCRSRRTTS---TVASATTPTQARRSPPPG--PGRSRSPVRPTSPS--- 96

Qy 94 DGSLPLPVAGLAGAPQTPCVRGAIWRSIPLRAS--PTSNGTYRSA----PLLE 144  
Db 97 -----AASPNSP-TARARYFR-SVARATTPTTWAASSTMPPSPPREST 143

Qy 145 ALPQPWNRAASGFVKTATM 162  
Db 144 GAPGWRSSST--STTIM 158

---

RESULT 4  
US-09-252-991A-16906  
; Sequence 16906; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16906  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16906

Query Match 11.4%; Score 99.5; DB 4; Length 263;  
Best Local Similarity 26.9%; Pred. No. 0.019; Mismatches 7; Indels 65; Gaps 7;

Qy 14 PTVAHTSSSVAVRSLVEFTCRAGLDWVCCRERLPSGRNLLEVDSLSPR 72  
Db 112 PARASTSDSPSPGR-----VRRNPAVVA----PGSRLRLDPVPASP-----A 149

Qy 73 GPGLSPTGLGSM-----AMRAAGGDGSCL-----PVALGLAGAPQTPGVCRAI 117  
Db 150 GFGSSGSPGPAGATGHRSRSSGERSAQCADRFFSGQRPAADDGPGAPARTGRGRR 209

Qy 118 WVRSSPLRLAASPTSMGTYRSSAPLLEALPOPWMA 153  
Db 210 WV-----PGPWRCM 218

---

RESULT 5  
US-09-252-991A-18207  
; Sequence 18207; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18207

Query Match 11.4%; Score 98; DB 4; Length 162;  
Best Local Similarity 25.3%; Pred. No. 0.015; Mismatches 24; Indels 30; Gaps 8;

Qy 15 TVAHRPSSSVAVRSLVEFTCRAGLDWVCCRERLPSGRNLLEVDSLSPR 75  
Db 6 SVPRTPASPAISSRRI-ACCSATCSSLACARKDAWPYSPSTAWRAAKATRISLCHW 63

Qy 60 DVSLSPRLVGRAGPGLSPGTLPSPMAMRAAGGRDGSCLPVALGLAGAPQ---TGVGR 115  
Db 64 GTAFSSRASSTTSFTPTTRPACNWRA-----TASCPASRSASGSPTAACSGRSP 120

Qy 116 AIWRSIPLRAASPTSWGTYRSSAASPTLGPGRMASFWKTAT 161  
Db 121 ATAAKSATSTAAATSPCRHRSTSN-GSGIPARW-----WRAAS 160

---

RESULT 6  
US-09-252-991A-24790  
; Sequence 24790; Application US/09252991A  
; Patent No. 6551795

GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO: 24790  
 LENGTH: 202  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-24790

Query Match 11.4%; Score 98; DB 4; Length 202;  
 Best Local Similarity 25.1%; Pred. No. 0.02;  
 Matches 46; Conservative 12; Mismatches 71; Indels 54; Gaps 7;

Qy 14 PTVAHRTSSRRAVAVSLVEFTCCRAGALDWCA-----RRERLPSGNANLEVDVSSLSSR 66  
 Db 1 PSTMSMSPRLPVRP-----CSTCSMPMACAGSRSATMARPAAVAYATAAWS 54

Qy 67 LVDPRAGPGL-----SPG-----  
 Db 55 TAGPSAAPRSPSSGRACGWKPKAATSTRNGHEPAGDERRLGHGRSHARGPR 98

Qy 99 PVALGIA---GAPOTPGVGRAIWVRSSIPILRAASPTSMGT-----YRSSAPPLEALFG 148  
 Db 115 PLRREGIASRQLSRPLAAGR----RSGLPRLHDAIGKTAARFLRLRAHRPAEQPQPG 171

Qy 149 PWR 151  
 Db 172 RWR 174

RESULT 7  
 US-09-252-991A-22801  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 SEQ ID NO: 22801  
 LENGTH: 231  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-22801

Query Match 11.3%; Score 97; DB 4; Length 231;  
 Best Local Similarity 29.6%; Pred. No. 0.03;  
 Matches 45; Conservative 8; Mismatches 55; Indels 44; Gaps 7;

Qy 10 TNVTPVAVRHTSSRRAVAVSLVEFTCCRAGALDW-YCARRERPSGRNLLEVDSLSPRL 67  
 Db 8 TRTAGACSSRCSSAR-----CHSATRMPGPCARSALPSGR-----RG 47

Qy 68 VGPRAAGPGLSPGTLGPSMAMRAAGGRDGSCLPVAGLQAGAPQTPGVGRAIWVRSSIPURA 127  
 Db 48 SGSTAGTNWPSPATTSMSTACIARSRT---TIANIRRAPMP-----PLTV 91

Qy 120 ASPS---WGTYRSSAPPLEALGPWNMASG 155

Db 92 SSPTSSRRPWIPH-SAMTREAIRAWRIASG 121  
 RESULT 8  
 US-09-252-991A-27305  
 Sequence 27305, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO: 27305  
 LENGTH: 709  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-27305

Query Match 11.2%; Score 96; DB 4; Length 709;  
 Best Local Similarity 34.5%; Prod. No. 0.16;  
 Matches 50; Conservative 6; Mismatches 65; Indels 24; Gaps 9;

Qy 19 RTSSSRVARSLVEFTCCRAGALDWCARRELPSGRNLLEVDSLSP-----RLVGR 71  
 Db 96 RTPGGRPACRV-----TRARRSAL-ADAGRGGLPAGQRQRPDALAPRPGAGERLSRSR 150

Qy 72 AGPCLSPGTLGPSMAMRAAGGRDGSCLPVAGLQAGAPQTPGVGRAIWVRSSIPURA 131  
 Db 151 AVAALRPES-RPRLAREAGGAGGA--PPEPGVALEPDRGAGR-WLPGG-AFPAAFPG 204

Qy 132 SWGTYRSSAPPLFALGPFWMASGF 156  
 Db 205 RLG----SRP--GAQPGFGLRACGF 223

RESULT 9  
 US-09-252-991A-17729  
 Sequence 17729, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO: 17729  
 LENGTH: 328  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-17729

Query Match 11.1%; Score 95.5%; DB 4; Length 328;  
 Best Local Similarity 29.4%; Prod. No. 0.067;  
 Matches 48; Conservative 12; Mismatches 60; Indels 43; Gaps 8;

Qy 14 PTVAHRTSSRRAVAVSLVEFTCCRAGALDWCARRELPSGR-----NIEVDV 61  
 Db 95 PVPVHHRPRSSR-----PHGA-----AR-----PAGRRSGCQDHGRGPDPHSPV 133

Qy 62 SLSPRLVGPRAGPGLSPGTL-GPSMAMRAACG-RDG-----CLPVALGLGAGQTPG 112

RESULT 10  
 US-09-252-991A-25006  
 ; Sequence 2506, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 25006  
 ; LENGTH: 371  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-25006

Query Match 11.1%; Score 95; DB 4; Length 371;  
 Best Local Similarity 32.6%; Pred. No. 0.080;  
 Matches 44; Conservative 10; Mismatches 53; Indels 28; Gaps 6;

Qy 47 RERDPSGMNLDEVDSLSPLVPGPAGPGLSPGTLGPMAMRAAGGRDGSCLPVALGAG 106  
 Db 31 RRDPGPWVRLQVADRPRAAPAGRG-APG-----RORAGRRTARPVAAAGAG 82

Qy 107 ---AQQTGIVGTAIIVRSI----PLR---AASPTSW-----GYRSSAPLLEA 146  
 Db 83 DSWAHRHGYFGDAQDLAQSTAHRPPGRGPPAPSPAWRAAAGTGTTARTGRPRAAT 142

Qy 147 PGPNRMASGFWKTTAT 161  
 Db 143 PARGUSAPALRRPAT 157

RESULT 11  
 US-09-252-991A-24258  
 ; Sequence 24258, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 24258  
 ; LENGTH: 213  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-24258

Query Match 10.9%; Score 93; DB 4; Length 213;  
 Best Local Similarity 29.8%; Pred. No. 0.07;  
 Matches 54; Conservative 13; Mismatches 60; Indels 54; Gaps 10;

RESULT 12  
 US-09-252-991A-23311  
 ; Sequence 23311, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 23311  
 ; LENGTH: 254  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-23311

Query Match 10.8%; Score 92.5%; DB 4; Length 254;  
 Best Local Similarity 29.6%; Pred. No. 0.099;  
 Matches 50; Conservative 12; Mismatches 80; Indels 27; Gaps 8;

Qy 13 TPTVAHRTSSSRVAVRSIVEFTCC-----RAGALDWVCAERRLPLS-GRNLEYDV 61  
 Db 86 TPAPARGRADERLVRSPARLALPGSAAROARRSGATLQCARPRQAQRRIIRPGS 145

Qy 62 SUSP-----RLVGRPAGPGLSPGTLSPTGPMAMRAAGGRDGSCLPVALGAGP-QTPCGV 114  
 Db 146 ATAPGTAPEFRRLQPRPAAHRWPSPAPVAGHPGRPEPLPGAGRHRHPPRVQACPGV 205

Qy 115 RAIWRSSTIP-----RAASP-TSIVOTYRSAPLALPG-PWRMASG 155  
 Db 206 -ATGLRADPGAGYGTILRRGRPEPGGAARRGTPGLELRPGSPAETARG 253

RESULT 13  
 US-09-252-991A-24717  
 ; Sequence 24717, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 24717

Page 5

LENGTH: 461  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-24717  
 Query Match 10.7%; Score 91.5; DB 4; Length 461;  
 Best Local Similarity 28.2%; Pred. No. 0.27; Indels 39; Gaps 7;  
 Matches 42; Conservative 14; Mismatches 54; Indels 39; Gaps 7;  
 SEQ ID NO: 26216 ; LENGTH: 1027  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-26216  
 Qy 44 VCARRELESGR----NLEFDVSLSPR----LVG----PRAGP----GL----76  
 Db 228 VAGARRLPQKLLGAALRLRRPSRSRPLRRCRGLAGNGLATYPAAGPGVGLPAC 287  
 Qy 77 -----SPGTGLPSSMAMRANGRDGSCLPVAGLLAGAPQPQPGVGRAIWRSIPLR 126  
 Db 288 LRYTRDRLRPPVAALPEVARSQGCACTA-PAGGGAGGKSPAGLGR-LARRSFPSS 345  
 Qy 127 AASPTSWGTYRSSAPLLEALPGPWRMASG 155  
 Db 346 SATPPAATTYFAARAPMPAITEPGRSGTG 374  
 RESULT 14  
 US-09-252-991A-29632  
 ; Sequence 29632, Application US/09252991A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO: 29632  
 ; LENGTH: 253  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-29632  
 Query Match 10.6%; Score 91; DB 4; Length 253;  
 Best Local Similarity 32.6%; Pred. No. 0.14; Indels 41; Mismatches 10; Gaps 8;  
 Matches 43; Conservative 10; Indels 38; Mismatches 41; Gaps 8;  
 Qy 46 ARERELPGRNLVED-----VSLSPRLVGPAGPGLSPGTLGPS-----NAMR 88  
 Db 131 ARRSPGPAGRVRGAAGATGPPAVRGPVRYGRATSSAPGRRAGR 190  
 Qy 89 AAGdGRDGSCLP----VALGLLAGAPQT-PGVGRAIWRSIPLRAS-PTSW----GTYRS 138  
 Db 191 PQRREGSGPRGRGRERRGRSSAPRSGRGTGRA-----CRAAGRGTSWRRLGTGRC 242  
 Qy 139 SAPPLEALPGPW 150  
 Db 243 RSP---ASPGGW 251  
 RESULT 15  
 US-09-252-991A-26216  
 ; Sequence 26216, Application US/09252991A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; LENGTH: 1027  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-26216  
 Qy 46 ARERELPGRNLVED-----VSLSPRLVGPAGPGLSPGTLGPS-----NAMR 88  
 Db 131 ARRSPGPAGRVRGAAGATGPPAVRGPVRYGRATSSAPGRRAGR 190  
 Qy 89 AAGdGRDGSCLP----VALGLLAGAPQT-PGVGRAIWRSIPLRAS-PTSW----GTYRS 138  
 Db 191 PQRREGSGPRGRGRERRGRSSAPRSGRGTGRA-----CRAAGRGTSWRRLGTGRC 242  
 Qy 139 SAPPLEALPGPW 150  
 Db 243 RSP---ASPGGW 251

Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw mode!

Run on: August 25, 2005, 03:08:37 ; Search time 161 Seconds  
(without alignments)

Scoring table: BLOSUM62  
Gapext 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing First 45 summaries:

Database : Published Applications AA:  
 1: /cgn2\_6/\_ptodata/1/pubpa/US07\_PUBCOMB.pep:  
 2: /cgn2\_6/\_ptodata/1/pubpa/PCT\_NEW\_PUB.pep:  
 3: /cgn2\_6/\_ptodata/1/pubpa/US06\_PUB.pep:  
 4: /cgn2\_6/\_ptodata/1/pubpa/US05\_PUBCOMB.pep:  
 5: /cgn2\_6/\_ptodata/1/pubpa/US07\_NEW\_PUB.pep:  
 6: /cgn2\_6/\_ptodata/1/pubpa/PECTUS\_PUBCOMB.pep:  
 7: /cgn2\_6/\_ptodata/1/pubpa/US08\_NEW\_PUB.pep:  
 8: /cgn2\_6/\_ptodata/1/pubpa/US08\_PUBCOMB.pep:  
 9: /cgn2\_6/\_ptodata/1/pubpa/US09\_PUBCOMB.pep:  
 10: /cgn2\_6/\_ptodata/1/pubpa/US09\_PUBCOMB.pep:  
 11: /cgn2\_6/\_ptodata/1/pubpa/US09\_PUBCOMB.pep:  
 12: /cgn2\_6/\_ptodata/1/pubpa/US09\_NEW\_PUB.pep:  
 13: /cgn2\_6/\_ptodata/1/pubpa/US10A\_PUBCOMB.pep:  
 14: /cgn2\_6/\_ptodata/1/pubpa/US10B\_PUBCOMB.pep:  
 15: /cgn2\_6/\_ptodata/1/pubpa/US10C\_PUBCOMB.pep:  
 16: /cgn2\_6/\_ptodata/1/pubpa/US10D\_PUBCOMB.pep:  
 17: /cgn2\_6/\_ptodata/1/pubpa/US10E\_PUBCOMB.pep:  
 18: /cgn2\_6/\_ptodata/1/pubpa/US10F\_PUBCOMB.pep:  
 19: /cgn2\_6/\_ptodata/1/pubpa/US11A\_PUBCOMB.pep:  
 20: /cgn2\_6/\_ptodata/1/pubpa/US11B\_PUBCOMB.pep:  
 21: /cgn2\_6/\_ptodata/1/pubpa/US60\_NEW\_PUB.pep:  
 22: /cgn2\_6/\_ptodata/1/pubpa/US60\_PUBCOMB.pep:  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	857	100.0	164	17 US-10-664-038-1
2	837	97.7	160	15 US-10-150-233-1
3	824	96.1	196	16 US-10-601-020-2
4	824	96.1	196	16 US-10-601-020-10
5	806	94.0	161	9 US-09-736-959A-1
6	806	94.0	161	9 US-09-736-959A-32
7	768	89.6	161	9 US-09-736-959A-3
8	763	89.0	161	9 US-09-736-959A-2
9	658	76.8	192	16 US-10-601-020-17
10	648.5	75.7	195	16 US-10-601-020-9
11	624	72.8	161	9 US-09-736-959A-7

## ALIGNMENTS

RESULT 1  
US-10-664-038-1  
 / Sequence 1, Application US/10664038  
 ; Publication No. US2005005395A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VARAKLIOTI, AGORITIS  
 ; APPLICANT: GEORGOPOLOU, UTRANIA  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND NEW POLYPEPTIDES ASSOCIATED WITH  
 ; AND/OR OVERLAPPING WITH HEPATITIS C VIRUS CORE GENE  
 ; TITLE OF INVENTION: PRODUCTS  
 ; FILE REFERENCE: 03495-0194-00000  
 ; CURRENT APPLICATION NUMBER: US/10-664-038  
 ; CURRENT FILING DATE: 2003-07-17  
 ; PRIOR APPLICATION NUMBER: US/09-644-987  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: 60/151,074  
 ; PRIOR FILING DATE: 1999-08-27  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 1  
 ; LENGTH: 164  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis C virus  
 ; US-10-664-038-1

Query Match Score 857; DB 17; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-70;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARILNLKCKTNVTPVAHRTSSRVAVRSLVFTCCRAGLDWVCARRERLPSGRNLEVD 60  
 Db 1 ARILNLKCKTNVTPVAHRTSSRVAVRSLVFTCCRAGLDWVCARRERLPSGRNLEVD 60

RESULT 3  
 US-10-601-020-2  
 ; Sequence 2, Application US/10601020  
 ; Publication No. US20040156862A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Branch, Andrea D.  
 ; APPLICANT: Stump, Decherd D.  
 ; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 ; FILE REFERENCE: RII-003CPUSCN  
 ; CURRENT APPLICATION NUMBER: US/10/601,020  
 ; CURRENT FILING DATE: 2003-06-20  
 ; NUMBER OF SEQ ID NOS: 2  
 ; LENGTH: 160  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis C virus  
 ; FEATURE: P protein  
 ; OTHER INFORMATION: P protein  
 US-10-150-283-1  
 Query Match 97.7%; Score 837; DB 15; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 1.e-68;  
 Matches 160; Conservative 0; Mismatches 0; Gaps 0;  
 Db 1 ARILNLKKTNVTPYAHRTSSRVAVRSLVEFTCCRAGLDWVCARRRLPSGRNLEVD 60  
 1 ARILNLKKTNVTPYAHRTSSRVAVRSLVEFTCCRAGLDWVCARRRLPSGRNLEVD 60  
 Qy 61 VSLSPRLVGPRAAGPGLSPGTIGPSMAMRAAGGRDGSCLPYALGLAGAAPQTPGYGRAIWR 120  
 61 VSLSPRLVGPRAAGPGLSPGTIGPSMAMRAAGGRDGSCLPYALGLAGAAPQTPGYGRAIWR 120  
 Db 121 SSIPRAASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 160  
 121 SSIPRAASPTSWGTYRSAPLLEALPGPWRMASGFWKTA 160  
 RESULT 4  
 US-10-601-020-10  
 ; Sequence 10, Application US/10601020  
 ; Publication No. US20040156862A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Branch, Andrea D.  
 ; APPLICANT: Walewski, Jose L.  
 ; APPLICANT: Stump, Decherd D.  
 ; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
 ; FILE REFERENCE: RII-003CPUSCN  
 ; CURRENT APPLICATION NUMBER: US/10/601,020  
 ; CURRENT FILING DATE: 2003-06-20  
 ; PRIORITY APPLICATION NUMBER: US 09/719277  
 ; PRIORITY FILING DATE: 2001-04-13  
 ; PRIORITY APPLICATION NUMBER: US 60/088670  
 ; PRIORITY FILING DATE: 1998-06-09  
 ; PRIORITY APPLICATION NUMBER: US 60/089138  
 ; PRIORITY FILING DATE: 1998-06-11  
 ; PRIORITY APPLICATION NUMBER: PCT/US99/12929  
 ; PRIORITY FILING DATE: 1999-06-09  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 10  
 ; LENGTH: 196  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis C virus  
 US-10-601-020-10  
 Query Match 96.1%; Score 824; DB 16; Length 196;  
 Best Local Similarity 96.3%; Pred. No. 2.2e-67;  
 Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 5  
 US-09-736-959A-1  
 ; Sequence 1, Application US/09736959A  
 ; Publication No. US200007641A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of Southern California  
 ; APPLICANT: OU, Jing-Haiung

APPLICANT: XU, Zhenming  
TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS

FILE REFERENCE: 13761-742

CURRENT APPLICATION NUMBER: US/09/736,959A

CURRENT FILING DATE: 2000-12-14

PRIOR APPLICATION NUMBER: US 60/170,835

PRIOR FILING DATE: 1999-12-14

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 1

LENGTH: 161

TYPE: PRT

ORGANISM: Hepatitis C Virus

FEATURE:

OTHER INFORMATION: Exemplary P17 sequences

US-09-736-959A-1

Query Match

94.0%; Score 806; DB 9;

Length 161;

Best Local Similarity 99.4%;

Pred. No. 7.7e-66;

Mismatches 1;

Conservative 0;

Indels 0;

Gaps 0;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

FILE REFERENCE: 1376-959A-3

SEQUENCE 3, Application US/09/736,959A

PATENT NO. US20020076415A1

GENERAL INFORMATION:

APPLICANT: University of Southern California

APPLICANT: OU, Jing-Hsiung

APPLICANT: XU, Zhenming

TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS

FILE REFERENCE: 13761-742

CURRENT APPLICATION NUMBER: US/09/736,959A

CURRENT FILING DATE: 2000-12-14

PRIOR APPLICATION NUMBER: US 60/170,835

PRIOR FILING DATE: 1999-12-14

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 32

LENGTH: 161

TYPE: PRT

ORGANISM: Hepatitis C Virus (HCV-1 subtype)

FEATURE:

OTHER INFORMATION: Predicted P17 sequence

US-09-736-959A-32

Query Match

94.0%; Score 806; DB 9;

Length 161;

Best Local Similarity 99.4%;

Pred. No. 7.7e-66;

Mismatches 1;

Conservative 0;

Indels 0;

Gaps 0;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

RESULT 7

US-09-736-959A-3

SEQUENCE 3, Application US/09/736,959A

PATENT NO. US20020076415A1

GENERAL INFORMATION:

APPLICANT: University of Southern California

APPLICANT: OU, Jing-Hsiung

APPLICANT: XU, Zhenming

TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS

FILE REFERENCE: 13761-742

CURRENT APPLICATION NUMBER: US/09/736,959A

CURRENT FILING DATE: 2000-12-14

PRIOR APPLICATION NUMBER: US 60/170,835

PRIOR FILING DATE: 1999-12-14

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 1

LENGTH: 161

TYPE: PRT

ORGANISM: Hepatitis C Virus

FEATURE:

OTHER INFORMATION: Exemplary P17 sequences

US-09-736-959A-3

Query Match

89.6%; Score 768;

DB 9; Length 161;

Best Local Similarity 95.5%;

Pred. No. 2.3e-62;

Mismatches 5;

Indels 0;

Gaps 0;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

FILE REFERENCE: 1376-959A-2

SEQUENCE 2, Application US/09/736,959A

PATENT NO. US20020076415A1

GENERAL INFORMATION:

APPLICANT: University of Southern California

APPLICANT: OU, Jing-Hsiung

APPLICANT: XU, Zhenming

TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS

FILE REFERENCE: 13761-742

CURRENT APPLICATION NUMBER: US/09/736,959A

CURRENT FILING DATE: 2000-12-14

PRIOR APPLICATION NUMBER: US 60/170,835

PRIOR FILING DATE: 1999-12-14

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 2

LENGTH: 161

TYPE: PRT

ORGANISM: Hepatitis C Virus

FEATURE:

OTHER INFORMATION: Exemplary P17 sequences

US-09-736-959A-2

Query Match

94.0%; Score 763;

DB 9; Length 161;

Best Local Similarity 94.0%;

Pred. No. 6.7e-62;

Mismatches 6;

Indels 0;

Gaps 0;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

**RESULT 9**

US-10-601-020-17

; Sequence 1, Application US/10601020  
; Publication No. US20040156862A1

; GENERAL INFORMATION:  
; APPLICANT: Branch, Andrea D.  
; APPLICANT: Stump, DeCherd D.

; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
; FILE REFERENCE: RII-003CPUSCN

; CURRENT APPLICATION NUMBER: US/10/601,020  
; CURRENT FILING DATE: 2003-06-20  
; PRIOR APPLICATION NUMBER: US 09/719277  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/088670  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: US 60/089138  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: PCT/US99/12929  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 17  
; LENGTH: PRF  
; TYPE: PRF  
; ORGANISM: Hepatitis C virus  
; SOFTWARE: PatentIn Ver. 2.0

Query Match 76.8%; Score 658; DB 16; Length 192;  
Best Local Similarity 77.4%; Pred. No. 3.3e-52;  
Matches 127; Conservative 9; Indels 2; Gaps 1;

Qy 1 ARILNKKTNTPVAVHTSSRVAVRSLIVEFTCCRAGLDWVCAARRGLPSPGRNLVEVDLSLSPR 67  
Db 1 AQILNKEPNVTPTAAHRTSSRAVRSLIVEFTCCRAGAPGNTCARLGRLPSPGRNLVEG 60

Qy 61 VSLSPRLVPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVAVGLAGAPTPGVGRAIVWR 120  
Db 61 DNLSPRFAGPRAFPGLSPGTGLPSMAMRWGGQDGSCHGALGVGAPTPGVGRIVWR 120

Qy 121 SSIPRAASPTSMGTYRSAPLLEALPGPWRMASGFWKATMQ 164  
Db 121 SSIPHAGGSPTSMGTYRSAPLLEALPGPWRMASGFWRSTTQQ 162

**RESULT 10**

US-10-601-020-9

; Sequence 9, Application US/10601020  
; Publication No. US20040156862A1

; GENERAL INFORMATION:  
; APPLICANT: Branch, Andrea D.  
; APPLICANT: Stump, DeCherd D.  
; APPLICANT: Walewski, Jose L.

; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
; FILE REFERENCE: RII-003CPUSCN

; CURRENT APPLICATION NUMBER: US/10/601,020  
; CURRENT FILING DATE: 2003-06-20  
; PRIOR APPLICATION NUMBER: US 09/719277  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/088670  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: US 60/089138

Query Match 72.8%; Score 624; DB 9; Length 161;  
Best Local Similarity 77.3%; Pred. No. 3.5e-49;  
Matches 119; Conservative 7; Indels 28; Gaps 0;

Qy 7 KKTKTNTPVAVHTSSRVAVRSLIVEFTCCRAGLDWVCAARRGLPSPGRNLVEVDLSLSPR 66  
Db 8 QRKPNVTPTAAHRTSSRAVRSLIVEFTCCRAGAPGNTCARLGRLPSPGRNLVEGDNLSPR 67

Qy 67 LVGPRAGPGLSPGTGLPSMAMRAAGGRDGSCLPVAVGLAGAPTPGVGRAIVRSIPLR 126  
Db 68 LDPRAGPGLSPGTGLPSMAMRALGDGSCHGALGVGAPTPGVGRIVWRSS1PSH 127

Qy 127 ASPTSMGTYRSAPLLEALPGPWRMASGFWKATQA 160  
Db 128 ASPTSMGTYRSAPLLEALPGPWRMASGFWRHMSGFWRTA 161

RESULT 12

US-10-601-020-13

; Sequence 13, Application US/10601020  
; Publication No. US20040156862A1

GENERAL INFORMATION:  
; APPLICANT: Branch, Andrea D.  
; APPLICANT: Walewski, Jose L.  
; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF  
; FILE REFERENCE: RII-003CPUSCN  
; CURRENT APPLICATION NUMBER: US/10/601,020  
; CURRENT FILING DATE: 2003-06-20  
; PRIOR APPLICATION NUMBER: US 09/719277  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/0886670  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: US 60/089138  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: PCT/US99/12929  
; PRIOR FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
; US-10-601-020-13

Query Match 68.1%; Score 583.5; DB 16; Length 195;  
Best Local Similarity 73.8%; Pred. No. 2.2e-45;  
Matches 121; Conservative 2; Mismatches 40; Indels 1; Gaps 1;

Qy 1 ARILNLKKKNTNPTVAHRTSSSRVAVRSLLVEFTCCAGALDWVCCRRLPSGRNLEVD 60  
Db 1 ARILNLKKKNTNPTVAHRTSSSRVAVRSLLVEFTCCAGALDWVCCRRLPSGRNLEVD 60

Qy 61 VSDSPRLVGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWR 120  
Db 61 VSDSPRLVGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWR 120

Qy 121 SSPLRRAASPTNSGTYRSAPLLEALPGPWRMASGFNKATMQ 164  
Db 121 SSPLRRAASPTNSGTYRSAPLLEALPGPWRMASGFNKATMQ 164

RESULT 13  
US-09-736-959A-5  
; Sequence 5, Application US/09736959A  
; Patent No. US20020076415A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Southern California  
; APPLICANT: XU, Jing-Hsiung  
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS  
; FILE REFERENCE: 13761-742  
; CURRENT APPLICATION NUMBER: US/09/736,959A  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: US 60/170,835  
; PRIOR FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Hepatitis C Virus  
; FEATURE:  
; OTHER INFORMATION: Exemplary P17 sequences  
; US-09-736-959A-5

Query Match 64.4%; Score 552; DB 9; Length 143;  
Best Local Similarity 79.3%; Pred. No. 1.2e-42;  
Matches 107; Conservative 5; Mismatches 23; Indels 0; Gaps 0;

Qy 7 KKKTNTNPTVAHRTSSSRVAVRSLLVEFTCCAGALDWVCCRRLPSGRNLEVD 66  
Db 8 QRKPNVTPTAHRTSSSRVAVRSLLVEFTCCAGALDWVCCRRLPSGRNLEVD 67

RESULT 14  
US-09-736-959A-9  
; Sequence 9, Application US/09736959A  
; Patent No. US20020076415A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Southern California  
; APPLICANT: OU, Jing-Hsiung  
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS  
; FILE REFERENCE: 13761-742  
; CURRENT APPLICATION NUMBER: US/09/736,959A  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: US 60/170,835  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Hepatitis C Virus  
; FEATURE:  
; OTHER INFORMATION: Exemplary P17 sequences  
; US-09-736-959A-8

Query Match 67 LVGPRAGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWRSSIPSLR 126  
Db 68 LAGPRAGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWRSSIPSLR 127

Qy 127 AASPTNSGTYRSAP 141  
Db 128 AASPTNSGTYRSAP 142

Query Match 67 LVGPRAGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWRSSIPSLR 66  
Db 68 LADPRAGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWRSSIPSLR 67

RESULT 15  
US-09-736-959A-8  
; Sequence 8, Application US/09736959A  
; Patent No. US20020076415A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Southern California  
; APPLICANT: OU, Jing-Hsiung  
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS  
; FILE REFERENCE: 13761-742  
; CURRENT APPLICATION NUMBER: US/09/736,959A  
; CURRENT FILING DATE: 2000-12-14  
; PRIOR APPLICATION NUMBER: US 60/170,835  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Hepatitis C Virus  
; FEATURE:  
; OTHER INFORMATION: Exemplary P17 sequences  
; US-09-736-959A-8

Query Match 67 LVGPRAGPGLSPGTGGSMSMANRAAGGRDGSCLPVALGLAGAPQTPEGVRAIWRSSIPSLR 66  
Db 68 QRKPNVTPTAHRTSSSRVAVRSLLVEFTCCAGALDWVCCRRLPSGRNLEVD 67

Query Match      Score 547; DB 9; Length 143;  
Best Local Similarity      63.8%; Pred. No. 3; 5e-42;  
Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
Matches 106;保守性 6; 错配 23; 缺失 0; 插入 0;

---

Qy      7 KKKCTNVTPTVAHRTSSSRVAVRSILVEFTCCRAGALDWYCARRERLPSGRNLLEVDSLSPR 66  
Db      8 QRKSNVTPTAAHRTSSRAVRSILVEFTCCRAGAPGWYCARLGRPLPSGRNLVEGDNLSPR 67

Qy      67 LVGRAGPGLSPGTLGPSNAMRAAGGRDSCULPVALGLAGAPQTPGVCRATWRSISPLR 126  
Db      68 LASPRAGPGLSPGTPGPSNAMRAWGQDGSCHPAAPGLVGAPKTPGVGRVIWRSISPH 127

Qy      127 AASPTSWGYRSSAP 141.  
Db      128 AASPTSWGYRSSAP 142

Search completed: August 25, 2005, 03:29:13  
Job time : 162 secs